

acid indicated in the features, the parentheses indicated that the heterogeneity was detected at or near to the 5'- or 3'-end of the HCV in the clone.

The sequence is derived from a composite HCV cDNA from HCV1, a prototypic HCV. The DNA sequence is based upon sequence information derived from a no. of HCV cDNA clones, which were isolated from a no. of HCV cDNA libraries, including the "c" library present in lambda gt11 (ATCC No.40394), and from human serum. The HCV cDNA clones were isolated by methods described in W09014436.

The clones from which the sequence was derived are 5'clonc32, b114a, 18q, aq40a, CA203a, CA290a, CA216a, p114a, CA167b, CA156e, CA84a, CA53a, K9-1 (also called K9 1), 26j, 12f, 14f, 11b, 7f, 7e, 8h, 33c, 43b, 37b, 35, 36f, 81, 32, 37b, 25c, 14c, 8f, 33f, 33q, 39c, 35f, 19q, 26q, 15c, b5a, 16jh, 6k and p131b.

Sequence 3011 AA:

xx

Query Match 0.58; Score 59; DB 13; Length 3011;
Best local Similarity 34.5%; pred No 4 4e+03;
Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps

Q:	1821	VLLP	AIKE	TYK	QIE	KWN	KNHC	PFMS	ILQ	1849
		:	:	:	:	:	:	:	:	
pb	1745	vi	av	av	av	av	av	av	av	1773

RESULT	200
AAR31621	
ID	AAR31621 standard; protein; 3011 AA.
XX	
XX	
AC	AAR31621;
XX	
XX	
DT	24-MAY-1993 (first entry)
XX	
XX	
DE	Hepatitis C virus (HCV) polyprotein.
XX	
XX	
KW	Hepatitis; liver disease; HCV1; monoclonal antibody; epitope;
KW	immobilised reagent; immunassay; diagnosis; detection; treatment;
KW	infection.

XX	Hepatitis C virus 1.	
XX	Key	Location/Qualifiers
XX	BT	Misc-difference 9 /label= ARG
XX	BT	Misc-difference 11 /label= THR
XX	BT	Misc-difference 176 /label= THR
XX	BT	Misc-difference 334 /label= VAL
XX	BT	Misc-difference 603 /label= ILE
XX	BT	Misc-difference 848 /label= (ASN)
XX	BT	Misc-difference 1114 /label= SER
XX	BT	Misc-difference 1117 /label= THR
XX	BT	Misc-difference 1276 /label= LEU
XX	BT	Misc-difference 1328 /label= (VAL)
XX	BT	Misc-difference 1454 /label= TYR
XX	BT	Misc-difference 1471 /label= (SER)
XX	BT	Misc-difference 1877 /label= (GLY)
XX	BT	Misc-difference 1948 /label= (HIS)
XX	BT	Misc-difference 1949 /label= (CYS)
XX	BT	Misc-difference 2021 /label= (VAL)
XX	BT	Misc-difference 2349 /label= (SER)
XX	BT	Misc-difference 2385 /label= (PHE)
XX	BT	Misc-difference 2486 /label= (ALA)
XX	BT	Misc-difference 2502 /label= (PHE)
XX	BT	Misc-difference 2690 /label= (GLY)
XX	BT	Misc-difference 2996 /label= (PRO)
XX	XX	W9205642-A.
XX	XX	20-FEB-1992.
XX	XX	12-AUG-1991; 91W0-US05728.
XX	XX	10-AUG-1990; 90US-0566209.
XX	XX	(CHIR-) CHIRON CORP.
XX	XX	Houghton M, Choo QL, Kuo G, Weiner AJ, Urdea MS, Irwine BD; Kolberg JA;
XX	XX	WPL: 1992-080094/10.
XX	XX	N-PSDB; AAQ21744.
XX	XX	Reagents for isolating, amplifying and detecting HCV polynucleotide(s) - used to monitor spread of blood-borne non-a, non-b hepatitis virus infection and screen blood samples for virus
XX	XX	dislosure; fig 1: 67pp; English.
XX	XX	Heterogeneities in cloned DNAs of HCV1 are indicated by the amino

OS Hepatitis C virus.
 PN EP644202-A.
 PU 22-MAR-1995.
 PF 14-DEC-1990; 90EP-0124241.
 PP 14-DEC-1990; 90EP-0124241.
 PR 14-DEC-1990; 90EP-0108611.
 XX (INR)) LRRCHERITICS NV.
 XX Deleys RJ, Maertens G, Pollet D, Van Heuverswyn H;
 UR WPI; 1995-116946/16.
 XX Synthetic antigens for the detection of hepatitis C virus
 PT antibodies comprise portions of the HCV peptide sequence, for
 PT use in screening blood and blood products
 PT disclosure; Fig 1; 51pp; English.
 PS AAP70290 is the composite hepatitis C virus (HCV) antigen
 CC from which the synthetic HCV antigens described in AAP70210, 270229
 CC were derived. These synthetic antigens can be used to screen blood,
 CC or blood products for the presence of HCV. They can also be used in
 CC various specific assays for the detection of HCV antibodies, and
 CC antigens, or as immunogens.
 XX Sequence 2894 AA;
 SQ

Query Match: 0.5%; Score 59; DB 16; Length 2894.
 Best Local Similarity 34.5%; Pred. No. 4 3a+03;
 Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0.

OY 1821 VLLPA:KKTYKQIEKNWKNMGPPMSILQ 1849
 DB 1745 viapavqlwqkklwtwaklmwfnlsiq 1773
 L: 11:: : : : : | | | | |

RESULT 196
 AAR08124
 ID AAR08124 standard; protein; 2955 AA.
 AC AAR08124;
 XX 23-JAN-1991 (first entry)
 DT Hepatitis C virus putative polyprotein.
 DE Hepatitis C virus (HCV), antiviral agent.
 XX Hepatitis C virus.
 OS Key Localization/Qualifiers
 EH Misc-difference 9...9 /label K or R
 FT Misc-difference 11...11 /label N or T
 FT Misc-difference 176...176 /label I or T
 FT Misc-difference 334...334 /label M or V
 FT Misc-difference 603...603 /label I or L
 FT Misc-difference 848...848 /label Y or N
 FT Misc-difference 1114...1114 /label P or S
 FT Misc-difference 1117...1117 /label S or T

PT Misc-difference 1276...1276 /label-P or L
 FT Misc-difference 1454...1454 /label=C or Y
 FT Misc-difference 1471...1471 /label-T or S
 FT Misc-difference 1877...1877 /label-E or G
 FT Misc-difference 1948...1948 /label-L or H
 FT Misc-difference 1949...1949 /label-S or C
 FT Misc-difference 2021...2021 /label-V or G
 FT Misc-difference 2349...2349 /label-T or S
 FT Misc-difference 2385...2385 /label-Y or F
 FT Misc-difference 2386...2386 /label-S or A
 FT Misc-difference 2502...2502 /label-L or F
 FT Misc-difference 2690...2690 /label-F or G
 FT Misc-difference 2921...2921 /label-R or G
 XX EP388232-A.
 PN 19-SEP-1990.
 XX 16-MAR-1990; 90EP-0302866.
 XX 18-MAY-1989; 89US-0355002.
 PP 17-MAP-1989; 89US-0355338.
 PR 20-APR-1989; 89US-0341334.
 XX (CHIP-) CHIPON CORP.
 PA Houghton M, Choo QL, Kuo G;
 PL WPI; 1990-284418/38.
 DR N-PSDB; AAQ05956.
 XX Hepatitis C virus DNA - used for producing probes,
 PT polypeptide(s), antibodies and anti-sense polynucleotide(s) for
 PT diagnosis and therapy.
 PS Disclosure; Fig 17; 83pp; English.
 XX HCV cDNA libraries were constructed using pooled serum from a
 CC chimpanzee with chronic HCV infection. A lambda gt11 library was
 CC screened with probes derived from previously isolated clones. The
 CC ORF is derived from the overlapping clones b114a, aq30a, CA205a,
 CC CA290a, CA210a, f14a, fALa-b, CA156a, CA147a, CA55a, fALa-26), 13i,
 CC 12f, 14i, 11b, 7f, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 34b, 25c,
 CC 14*, 81, 30i, 83i, 85i, 86i, 18g, 26g, 15e, b5a and 16jh.
 CC Polypeptide encoded by this sequence can be used to design probes
 CC for the detection of HCV nucleic acids, in screening programmes
 CC for antiviral agents and in preparing blood free of HCV. The
 CC sequence contains 189 (overlapping) peptides which are claimed as
 CC HCV epitopes.
 CC See also AAQ05955.
 XX Sequence 2955 AA;
 SQ

Query Match: 0.5%; Score 59; DB 11; Length 2955;
 Best Local Similarity 34.5%; Pred. No. 4 3a+03;
 Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

OY 1821 VLLPA:KKTYKQIEKNWKNMGPPMSILQ 1849
 L: 11:: : : : : | | | | |

[illegible]

```

XX Sequence: 2549 AA;
SQ
Query Match: 0.5%; Score 59; DB 17; Length 2549;
Best Local Similarity: 41.4%; Ident. No. 450004;
Matches: 11; Conservative 12; Mismatches 12; Indels 0; Gaps 0;

QY 2570 LKRTQSEPKVFEAAITVIAAKFKENIVILIP 2594
1159 ivrtldspelslstdmrlsslvrlqkkyqltlp 1194
1159 ivrtldspelslstdmrlsslvrlqkkyqltlp 1194

RESULT 191
AAW4245
XX AAW4245 standard; Protein: 2549 AA;
XX AAW4245;
XX 28-AUG-1998 (first entry)
XX FRAP (tor1) protein.
XX FRAP: tor1 protein; macrolide binding protein;
XX graft versus host disease; stem cell; gene therapy.
XX Mammalia.
XX W9808956 A2.
XX 05-MAR-1998.
XX 27-AUG-1997; 57W5 0315153.
XX 27-AUG-1996; 56US-0024666.
XX (HARD ) HARVARD COLLEGE.
XX (STRO ) UNIV LELAND STANFORD JUNIOR.
XX Bolshaw PJ, Crabtree G, Schreiber SL;
XX WPI: 1998-179441/16.
XX N-PSDB; AAV28518.
XX Selective inhibition of proliferation of haematopoietic cells
XX using macrolide binding proteins and antibodies, useful for treatment
XX of graft versus host disease
XX Disclosure: Page 64-74; 85pp; English.
XX This polypeptide comprises FRAP (tor1) protein. A claimed method
XX for selective inhibition of proliferation of a haematopoietic cell
XX comprises contacting a haematopoietic cell which ectopically
XX expresses a gene encoding a mutated macrolide binding protein
XX (Mbp), which has altered macrolide-binding specificity relative to
XX the wild-type term Mbp, with a macrolide which selectively induces
XX macrolide dependent inhibition of proliferation of cells expressing
XX the mutated Mbp compared to cells expressing a wild-type form of
XX the Mbp. The Mbp is selected from FRAP, an FK506 binding protein
XX (see AAV28517), a cyclophilin and a calcineurin (see AAV6628). Also
XX claimed are: an expression construct encoding a mutated Mbp chosen
XX from FRAP, FKBP, cyclophilin and calcineurin; a method for
XX selectively inhibiting proliferation of a transplanted
XX haematopoietic cell; a method for treating graft versus-host
XX disease by selectively inhibiting proliferation of transplanted
XX haematopoietic cells; and a method of promoting engraftment and
XX haematopoietic activity of a haematopoietic stem cell from a donor.
XX Sequence: 2549 AA;
SQ
Query Match: 0.5%; Score 59; DB 17; Length 2549;
Best Local Similarity: 41.4%; Ident. No. 450004;
Matches: 11; Conservative 12; Mismatches 12; Indels 0; Gaps 0;

QY 2570 LKRTQSEPKVFEAAITVIAAKFKENIVILIP 2594
1159 ivrtldspelslstdmrlsslvrlqkkyqltlp 1194
1159 ivrtldspelslstdmrlsslvrlqkkyqltlp 1194

RESULT 191
AAW4245
XX AAW4245 standard; protein: 2772 AA.
XX AAW4245;
XX 28-AUG-1998 (first entry)
XX FRAP (tor1) protein.
XX FRAP: tor1 protein; macrolide binding protein;
XX graft versus host disease; stem cell; gene therapy.
XX Mammalia.
XX W9808956 A2.
XX 05-MAR-1998.
XX 27-AUG-1997; 57W5 0315153.
XX 27-AUG-1996; 56US-0024666.
XX (HARD ) HARVARD COLLEGE.
XX (STRO ) UNIV LELAND STANFORD JUNIOR.
XX Bolshaw PJ, Crabtree G, Schreiber SL;
XX WPI: 1998-179441/16.
XX N-PSDB; AAV28518.
XX Selective inhibition of proliferation of haematopoietic cells
XX using macrolide binding proteins and antibodies, useful for treatment
XX of graft versus host disease
XX Disclosure: Page 64-74; 85pp; English.
XX This polypeptide comprises FRAP (tor1) protein. A claimed method
XX for selective inhibition of proliferation of a haematopoietic cell
XX comprises contacting a haematopoietic cell which ectopically
XX expresses a gene encoding a mutated macrolide binding protein
XX (Mbp), which has altered macrolide-binding specificity relative to
XX the wild-type term Mbp, with a macrolide which selectively induces
XX macrolide dependent inhibition of proliferation of cells expressing
XX the mutated Mbp compared to cells expressing a wild-type form of
XX the Mbp. The Mbp is selected from FRAP, an FK506 binding protein
XX (see AAV28517), a cyclophilin and a calcineurin (see AAV6628). Also
XX claimed are: an expression construct encoding a mutated Mbp chosen
XX from FRAP, FKBP, cyclophilin and calcineurin; a method for
XX selectively inhibiting proliferation of a transplanted
XX haematopoietic cell; a method for treating graft versus-host
XX disease by selectively inhibiting proliferation of transplanted
XX haematopoietic cells; and a method of promoting engraftment and
XX haematopoietic activity of a haematopoietic stem cell from a donor.
XX Sequence: 2772 AA;
SQ
Query Match: 0.5%; Score 59; DB 17; Length 2772;
Best Local Similarity: 41.4%; Ident. No. 450004;
Matches: 11; Conservative 12; Mismatches 12; Indels 0; Gaps 0;

QY 2772 LKRTQSEPKVFEAAITVIAAKFKENIVILIP 2796
1159 ivrtldspelslstdmrlsslvrlqkkyqltlp 1194
1159 ivrtldspelslstdmrlsslvrlqkkyqltlp 1194

```

/note- "clained peptide reactive to HCV antibody"

PT JP04288097-A.
 XX 13-OCT-1992.
 XX 07-NOV-1991; 91JP-0401705.
 XX 07-NOV-1991; 91JP 0301705.
 XX (CHAU) GENEKUS OPTICAL CO LTD.
 XX WPI; 1992 36721/47.
 XX New peptide(s) are reactive to an antibody against type C
 PT hepatitis virus, used to detect virus in patients
 XX Disclosure; Fig 1; 11pp; Japanese.
 XX The peptides (clained) can be used to detect hepatitis C virus (HCV)
 CC antibody positive patients, post-transfusion hepatitis can be prevented
 CC by screening a sample from a HCV antibody positive patient with the
 CC peptide.
 XX Sequence 2436 AA;
 SQ
 Query Match 0.5%; Score 59; DB 13; Length 2436;
 Best Local Similarity 34.5%; Pred. No. 3,36+03;
 Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;
 QY 1821 VLLPAIKKTYGQIEKRWKHNHMCDFMSHLQ 1849
 DB 1295 viapqrqtnwqkietfwakkmwntfsgiq 1323
 RESULT 187
 AAP90288
 ID AAP90288 standard; protein; 2462 AA.
 XX AC AAP90288;
 XX 01-NOV 1989 (first entry)
 XX Peptide encoded by composite hepatitis C cDNA.
 DE Hepatitis C virus; clone 15c; clone 14-1; probe; vaccine.
 FW Pan troglodytes.
 OS GB2212511-A.
 XX 26-JUL-1989.
 XX 18-NOV-1988; 88GB-0027024.
 XX 18-NOV-1987; 87US-0122714.
 XX (CHUR) CHURCH CORPORATION.
 XX Houghton M; Choo QL; Kuo G;
 XX WPI; 1989-215054/30.
 XX N-PSDB; AAN90336.
 XX Hepatitis C virus gene - used for produ. of polynucleotide probes,
 PT polypeptide(s) and antibodies for diagnosis, prevention and
 PT treatment of infection.
 XX Disclosure; fig 47-1 to 47-8; 245pp; English.
 XX The sequence is the peptide encoded by the composite hepatitis C
 CC virus (HCV) cDNA of AAN90336. The polypeptides are used to

CC diagnose HCV-induced NANBH, to raise antibodies for
 CC immunosay or treatment, or to produce vaccines.

SQ Sequence 2462 AA;

Query Match 0.5%; Score 59; DB 10; Length 2462;
 Best Local Similarity 34.5%; Pred. No. 3,36+03;
 Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1821 VLLPAIKKTYGQIEKRWKHNHMCDFMSHLQ 1849

DB 1321 viapqrqtnwqkietfwakkmwntfsgiq 1349

RESULT 188

AAR81730

ID AAR81730 standard; Protein; 2549 AA.

XX AC AAR81730;

XX 29 MAY-1996 (first entry)

XX Sirolimus effector protein.

XX Glutathione-S-transferase; GST; FK506 binding protein; FKBP; PKC;
 KW rapamycin; FKBP-rapamycin binding protein; Mol14 cell; amplity;
 EW fusion protein; GST-FKBP12; immunomodulatory agent; primer;
 KW antitumour agent; detection; antisense DNA; immune system.

XX Homo sapiens.

XX EP676471-A2.

XX 11-OCT-1995.

XX 07-MAR-1995; 95EP-0301475.

XX 13 FEB-1995; 95US 0304524

XX 08-MAR-1994; 94US-0207975.

XX 26-SEP-1994; 94US-0312023.

XX (AMHP) AMERICAN HOME PROD CORP.

XX (UYCO) UNIV COLUMBIA NEW YORK.

XX Caggiano TJ; Chen Y; Failli AA; Molnar-Kimber KL;

XX Nakanishi K;

XX WPI; 1995-346091/45.

XX N-PSDB; AAN100770.

XX New effector proteins of rapamycin - which bind to a
 PT glutathione S-transferase FK506 binding protein-rapamycin complex

XX Example 2; Page 16-29; 44pp; English.

XX This sequence represents the sirolimus effector protein (SEP). The
 CC cDNA encoding this sequence was isolated from Mol14 human T-cell
 CC leukemia cells (ATCC CCL 1582) and used in the production of a fusion
 CC protein between glutathione S transferase (GST) and SEP. The sep gene
 CC was amplified in live fragments using the primers given in AAT00756-69.
 CC The amplified SEP gene was inserted into a vector already containing
 CC the GST gene and the fusion protein expressed (see also AAR81731). The
 CC fusion protein was used in the isolation of a protein of mammalian
 CC origin which binds a glutathione-S-transferase (GST)-FK506 binding
 CC protein (FKBP) rapamycin complex. The FKBP-rapamycin binding protein
 CC and corresponding DNA was isolated from Mol14 cells using a complex of
 CC the fusion protein GST FKBP12 and rapamycin. The isolated proteins have
 CC molecular weights of 125, 148, 298 and 210 kD. They can be used for
 CC identifying an immunomodulatory, or an antitumour agent. They can also
 CC be used in the detection of rapamycin, rapamycin analogues or
 CC metabolites when complexed with FKBP. Antisense DNA can be used to
 CC modulate the immune system of a mammal.

[illegible]

Db 422 viapavqtnwqkietfwakmwnfslsqiq 450
 I: |||: |||: | | | | | | | |

RESULT 160

AAR21565
 ID AAR21565 standard; Protein: 781 AA.

XX AAR21565;

DE 09-JUN-1992 (first entry)

XX HCV CKS-33-BCD - pHCV-31.

XX Hepatitis C virus; antigen; diagnosis; inhibitor; CMP-KDO synthase;
 KW CKS; HCV CKS-33-BCD; c100-3; NANBHV.

XX Key Location/Qualifiers

PT Peptide 1..239

PT Region /Label: CKS

PT Peptide /Label: linker

PT Peptide 240..247

PT Peptide 248..513

PT Peptide /Label: 33

PT Region /Note: "HCV region NS-3, amino acids 1192-1457"

PT Region 514..515

PT Peptide /Label: linker

PT Peptide 516..771

PT Peptide /Label: BCD

PT Region /Note: "HCV region NS-3, amino acids 1676-1931"

PT Region 772..781

PT Peptide /Label: linker

XX EP472207-A.

PR 26-FEB-1992.

XX 23-AUG-1991; 91EP-0114161.

PR 07-NOV-1990; 90US-0614069.

PR 24-AUG-1993; 90US-0572822.

XX (ABHO) ABBOTT LABORATORIES.

PT Devare SG, Desai SM, Casey JM, Lawson JJ, Lesniowski FF,

PT Bailey SH, Gutierrez PA, Stewart JL,

XX WPI; 1992 066430/09.

DR N-PSDB; AA021678.

XX Recombinant hepatitis C virus antigens - produced as fusion

PT proteins and representing distinct antigenic regions of the HCV

PT genome

XX Disclosure; Fig 11, page 37-43; 115pp; English.

PS The protein (mol.wt. 90 kDa), encoded by pHCV-31, is composed of two

XX non-continuous coding regions located in the putative non-structural

CC regions of HCV designated NS-3 and NS-4. Clone BCD represents the

CC C-terminal 256 amino acids of c100-3; the N-terminal amino acids are

CC not represented.

CC The polypeptide represents a distinct antigenic region of the HCV

CC genome and can be used for the detection of antibodies and antigens

CC for early diagnosis of HCV infection. The polypeptide can also be

CC used to develop specific inhibitors of viral replication and for

CC therapeutic purposes.

XX Sequence 781 AA;

SQ

Query Match 0.5%; Score 59; DB 13; Length 781;

Best Local Similarity 34.5%; Pred. No. 6,2e+02;

Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1821 VLLPAIKTKTYQIFKWNFNHMGPMMSIIQ 1849
 I: |||: |||: | | | | | | | |
 Db 585 viapavqtnwqkietfwakmwnfslsqiq 613

RESULT 161

AAR22208
 ID AAR22208 standard; Protein: 781 AA.

XX AAR22208;

DE 23-JUL-1992 (first entry)

XX Sequence of fusion protein HCV CKS-33-BCD recombinant antigen
 DE encoded by pHCV-31.

XX CKS fusion protein; antigen.

XX Hepatitis.

OS EP475182-A.

PN 18-MAR-1992.

XX 26-AUG-1991; 91EP-0114296.

XX 24-AUG-1990; 90US-0573103.

XX (ABHO) ABBOTT LABORATORIES.

PL Rolling TJ, Mandecki W;

DR WPI; 1992-089971/12.

DR N-PSDB; AA022963.

XX Hepatitis C virus CKS fusion protein - produced using vector

PT contg. lac control region

XX Example; Fig 8, 40pp, English.

XX CKS methods of protein synthesis are known in the art, for inst. a

CC CKS method has been disclosed in EP A. 0331 961. Six individual

CC nucleotides representing AAS 11-150 of the HCV genome were ligated

CC together and cloned as a 456 base pair EcoRI-BamHI fragment into the

CC CKS fusion vector p10296 to produce pHCV 34. The resultant fusion

CC protein HCV CKS Core, consists of 239 AAs of CKS, 7 AAs contributed

CC by linker DNA sequences, and the first 150 AAs of HCV. To

CC construct the plasmid pHCV-31, the 781 base pair EcoRI-BamHI

CC fragment from pHCV-23 representing the HCV BCD region was

CC linker-adapted and ligated into pHCV-29. The resulting plasmid,

CC designated pHCV-31, expresses HCV CKS-33-BCD recombinant antigen

CC which consists of 239 AAs of CKS, 8 AAs contributed by linker DNA

CC sequences, 266 AAs of the HCV NS3 region (AAS 1192-1457), 2 AAs

CC contributed by linker DNA sequences, 256 AAs of the HCV NS4 region

CC (AAS 1676-1931), and 10 additional AAs contributed by linker DNA

CC sequences.

XX Sequence 781 AA;

SQ

Query Match 0.5%; Score 59; DB 13; Length 781;

Best Local Similarity 34.5%; Pred. No. 6,2e+02;

Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1821 VLLPAIKTKTYQIFKWNFNHMGPMMSIIQ 1849

I: |||: |||: | | | | | | | |

Db 585 viapavqtnwqkietfwakmwnfslsqiq 613

RESULT 162

AAR33632
 ID AAR33632 standard; Protein: 781 AA.

DT	01-JUN-1993	(first entry)
XX		
DE	HCV-C100D2 recombinant antigen encoded by pHCV 58.	
XX		
KW	Hepatitis C virus; NS3, c100 antigen, CKS fusion protein;	
KW	CMP-KDO synthetase; immunodot assay; Non-A, non-B hepatitis.	
XX		
SS	Hepatitis C virus.	
XX		
PN	W69304087-A.	
XX		
DD	04-MAR-1993.	
XX		
PP	Z1-AUG-1992; 92MO-US07187.	
XX		
PP	Z1-AUG-1991; 91US-0748566.	
XX		
PA	(Abbott) ABB6/TI LAB.	
XX		
PT	Casey JM, Desai SM, Pature SC, Kapprecht KP;	
XX		
DR	WPI: 1993-094940/TI.	
XX		
DR	N-PSDB: AAQ18243.	
XX		
PT	Hepatitis C assay using recombinant C-100 region antigens - for	
PT	detecting antibodies and antigen in body fluids from individuals	
XX	exposed to hepatitis C virus	
XX		
FE	Example 2, Page 119 121, 206ff., English.	
XX		
CC	The clone pHCV-58 was constructed to overcome poor expression	
XX	levels of the HCV CKS c100 recombinant antigen. HCV CKS-c100	
CC	antigen consists of 239 amino acids of CKS, eight amino acids	
CC	contributed by linker DNA sequences, 363 amino acids from HCV NS3	
CC	region (amino acids 1569-1931) and 10 additional amino acids	
CC	contributed by linker DNA sequences. A 63bp NlaIV-HaeIII fragment	
CC	was deleted from pHCV-24 (a plasmid which expresses the HCV CKS-	
CC	c100 antigen under control of the lac promoter). The pHCV-58	
CC	fusion protein (containing a 21 amino acid deletion, i.e. of HCV	
CC	amino acids 1609-1620) was expressed at a significantly higher	
CC	level than the pHCV-24 fusion protein.	
XX		
SQ	Sequence 599 AA:	
	Query Match 0.5%; Score 59; DB 14; Length 599;	
	Best Local Similarity 34.5%; Pred. No. 4,3e+02;	
	Matches 10; Conservative 9; Mismatches 10; Indels 0;	
DY	1921 VLIPAEKVEYFAISFNSNFWNMWPEWSHIC 1247	
	I I I I I I I I I I I I I I I I	
Db	403 viapavqtawqkletfwakhwntisig 431	
RESULT 156		
ID	AAR33601	
AC	AAR33601 standard; Protein: 599 AA.	
AC	AAR33601;	
XX		
DT	05-JUN-1993 (first entry)	
XX		
DE	HCV C100b2 recombinant antigen encoded by pHCV-58.	
XX		
KW	Hepatitis C Virus; non-A, non-B hepatitis virus; NANBH;	
KW	non-structural protein, CMP KDO synthetase; CKS fusion protein;	
KW	anti-CMP KDO synthetase dot blot assay; anti-CMP KDO	
XX	immunassay, pHCV 58.	
XX		
OS	Hepatitis C Virus.	
XX		
PN	W69304088-A.	
XX		

```

04-MAR-1993.
XX
XX
21-AUG-1992; 92WO-US07188.
XX
XX
21-AUG-1991; 9JNS-0748561
XX
XX
(ABBO ) ABBOTT LAB.
XX
XX
Dailey SH, Besai SM, Devare SG;
XX
XX
WPI: 1993-093941/11.
XX
XX
N-PSDB; AAQ38258.
XX
XX
Hepatitis C assay using recombinant NS1 region antigens - for
XX
XX
detecting antibodies and antigen in body fluids from individuals
XX
XX
exposed to hepatitis C virus
XX
XX
Example 3; Page 93-95; 175pp; English.
XX
XX
The clone pHCV-58 was constructed to overcome poor expression
XX
XX
levels of the HCV CRS-cl00 recombinant antigen. HCV CRS-cl00
XX
XX
antigen consists of 246 amino acids of CRS, eight amino acids
XX
XX
contributed by linker DNA sequences, 963 amino acids from pSV
XX
XX
region (amino acids 1869-1931) and 10 additional amino acids
XX
XX
contributed by linker DNA sequences. A 63bp NlaIV-HaeIII fragment
XX
XX
was deleted from pHCV-24 (a plasmid which expresses the HCV CRS-
XX
XX
cl00 antigen under control of the lac promoter). The pHCV-58
XX
XX
antigen protein (containing a 21 amino acid deletion, i.e. of HCV
XX
XX
amino acids 1640-1620) was expressed at a significantly higher
XX
XX
level than the pHCV-24 fusion protein.
XX
XX
Sequence 599 AA;
XX
XX
Query Match 0.53; Score 59; Pr 14; Length 599;
XX
XX
Best local Similarity 34.5%; Prcl. No. 4.3e-02;
XX
XX
Matches 10; Conservative 9; Mismatches 10; Indels 0;
XX
XX
QY 1821 VLLPAIKTKYQIEKNKNHMGPEMSIIQ 1849
XX
XX
DB 403 viapavqinwqkietfwakhhwnfisiq 431
XX
XX
RESULT 157
XX
XX
AAB51379
XX
XX
ID AAB51379 standard; Protein: 599 AA.
XX
XX
AC AAB51379;
XX
XX
DT
XX
XX
17-APR-2001 (first entry)
XX
XX
DE HCV recombinant antigen pHCV-58 protein sequence SEQ ID NO:18.
XX
XX
KW Hepatitis C virus; HCV: antigen; detection; antibody.
XX
XX
OS Hepatitis C virus.
XX
XX
PN US6172189-B1.
XX
XX
PP 09-JAN-2001.
XX
XX
PD
XX
XX
02-JUN-1997; 97US-0867611.
XX
XX
19-NOV-1992; 92US-0949843.
XX
XX
10-JAN-1994; 94US-0179896.
XX
XX
01-MAY-1996; 96US-0646757.
XX
XX
24-APR-1990; 90US-0572832.
XX
XX
07-NOV-1990; 90US-0614069.
XX
XX
21-AUG-1991; 91US-0748561.
XX
XX
21-AUG-1991; 91US-0748566.
XX
XX
29-OCT-1991; 91US-0748565.
XX
XX
(ABBO ) ABBOTT LAB.

```


PT exposed to hepatitis C virus

XX Example 8, Page 114 116, 206pp, English.

XX The clone pHCV-57 was constructed to overcome poor expression levels of the HCV cKS-c100 recombinant antigen. HCV cKS-c100 antigen consists of 239 amino acids of cKS, eight amino acids contributed by linker DNA sequences, 363 amino acids from HCV NS4 region (amino acids 1569-1981) and 10 additional amino acids contributed by linker DNA sequences. A 69bp Bdel fragment was deleted from pHCV-24 (a plasmid which expresses the HCV cKS-c100 antigen under control of the lac promoter). The pHCV-57 fusion protein (containing a 23 amino acid deletion, i.e. of HCV amino acids 1575-1597) was expressed at a significantly higher level than the pHCV-24 fusion protein.

XX Sequence 597 AA;

Query Match 0.5%; Score 59; DB 14; Length 597;
Best Local Similarity 34.5%; Pred. No. 4.2e+02;
Matches 19; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1821 VLLDAIKKTYKQIEKNWENHMCPEMSILQ 1849
I: |||: ||||: | || ||: ||
Db 401 viapavqawqkierfwakhwofisq 429

RESULT 151

AAR34600

ID AAR34600 standard; Protein: 597 AA.

XX

AC AAR34600;

XX

DT 05-JUN-1998 (first entry)

XX

DE HCV C100B1 recombinant antigen encoded by pHCV-57.

XX

XX hepatitis C Virus; non A, non B hepatitis virus, HAVBIL, non structural protein, ORF 1/5; Synthesizer: CDS fusion protein; 239-288 - chlamy. maker. A 69bp Bdel fragment was deleted from pHCV-24 (a plasmid which expresses the HCV cKS-c100 antigen under control of the lac promoter).

XX

OS Hepatitis C Virus.

XX

PN W093404088-A.

XX

PD 04-MAR-1993.

XX

PF 21-AUG-1992; 92W0-0807188.

XX

PR 21-AUG-1991; 91US-0748561.

XX

PA (ABBO) ABBOTT LAB.

XX

PI Bailey SH, Desai SM, Devare SG;

XX

DR WPI: 1993-093941/11.

DR N-PSDB; AAO38257.

XX

XX Hepatitis C assay using recombinant NS3 region antigens for detecting antibodies and antigen in body fluids from individuals exposed to hepatitis C virus

XX Example 8; Page 88-90; 175pp; English.

XX The clone pHCV 57 was constructed to overcome poor expression levels of the HCV cKS-c100 recombinant antigen. HCV cKS-c100 antigen consists of 239 amino acids of cKS, eight amino acids contributed by linker DNA sequences, 363 amino acids from HCV NS4 region (amino acids 1569-1981) and 10 additional amino acids contributed by linker DNA sequences. A 69bp Bdel fragment was deleted from pHCV-24 (a plasmid which expresses the HCV cKS c100

CC antigen under control of the lac promoter). The pHCV 57 fusion protein (containing a 23 amino acid deletion, i.e. of HCV amino acids 1575-1597) was expressed at a significantly higher level than the pHCV-24 fusion protein.

XX Sequence 597 AA;

Query Match 0.5%; Score 59; DB 14; Length 597;
Best Local Similarity 34.5%; Pred. No. 4.2e+02;
Matches 19; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1821 VLLDAIKKTYKQIEKNWENHMCPEMSILQ 1849
I: |||: ||||: | || ||: ||
Db 401 viapavqawqkierfwakhwofisq 429

RESULT 152

AAB51378

ID AAB51378 standard; Protein: 597 AA.

XX

AC AAB51378;

XX

DT 17-APR 2001 (first entry)

XX

DE HCV recombinant antigen pHCV-57 protein sequence SEQ ID NO:16.

XX

FW Hepatitis C virus; HCV; antigen; detection antibody.

XX

OS Hepatitis C Virus.

XX

PN US6172189-B1.

XX

PD 09-JAN-2001.

XX

PF 02-JUN-1997; 97US-0867611.

XX

PR 19-NOV-1992; 92US-0989843.

PR 10-2AR-1994; 94US-0179896.

PR 01-MAY-1996; 96US-0646757.

PR 24-APR-1990; 90US-0572823.

PR 07-NOV-1990; 90US-0614069.

PR 21-AUG-1991; 91US-0748561.

PR 21-AUG-1991; 91US-0748566.

PR 29-OCT-1991; 91US-0748565.

XX

PA (ABBO) ABBOTT LAB.

XX

PI Devare SG, Desai SM, Casey JM, Bailey SH, Dawson GL, Gutierrez RA; Lesniewski RP, Stewart JR, Rupprecht KK;

XX

DR WPI: 2001-122352/13.

DR N-PSDB; AAP32225.

XX

XX New recombinant antigens representing distinct antigenic regions of Hepatitis C virus (HCV) genome, useful for detection of antibodies and antigens in body fluids of individuals exposed to HCV

XX Example 8; Column 91-94; 167pp; English.

XX The present invention describes recombinant Hepatitis C virus (HCV) antigens (1). (1) is useful as a reagent for the detection of antibodies and antigen in body fluids from individuals exposed to HCV. The HCV assay uses reliable and efficient reagents and methods to accurately detect the presence of HCV antibodies in samples obtained from individuals suspected of having HCV infection. AAP32218 to AAP32235, AAB51371 to AAB51379 and AAB64001 to AAB64032 represent sequences used in the exemplification of the present invention.

XX Sequence 597 AA;

Query Match 0.5%; Score 59; DB 14; Length 597;

SQ Sequence 594 AA;

Query Match 0.5%; Score 59; DB 22; Length 594;

Best Local Similarity 34.5%; Pred. No. 4,26-62;

Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1821 VLLPAIKETVYKQIEKNNKNNHMPFMSHLQ 1849

L I I I I I I I I I I I I I I I I I I I

Db 403 viapavqtawqkietiwaklmwnflisq 431

RESULT 146

AAR65566

ID AAR65566 standard; protein; 594 AA.

XX

AC AAR65566;

XX

DT 01-JUL-1994 (first entry)

XX

DE CKS-HCV antigen fusion protein pHCV-63.

XX

KW Hepatitis C virus; C100 antigen; CKS fusion protein;

KW CMP RNA synthetase, immunodet assay, Neg A, neg B hepatitis

XX

GS Hepatitis C virus.

XX

PN W09404087-A.

XX

PD 04-MAR-1993.

XX

PF 21-AUG-1992; 92WS-US97187.

XX

PR 21-AUG-1991; 91US-0748566.

XX

PA (ABBO) ABBOTT LAB.

XX

PL Casey SM, Desai SM, Casey SM, Supprocht KP;

XX

XX WPI; 1993-0930409/11.

XX

PT Hepatitis C assay using recombinant C-100 region antigens for

PT detecting antibodies and antigen in body fluids from individuals

PT exposed to hepatitis C virus

XX

PS Claim 2; Page 55-57; 206pp; English.

XX

CC A specific antigenic region of the HCV genome is expressed as a

CC chimeric fusion with E. coli CMP RNA synthetase (CKS) gene. The

CC fusion protein pHCV-63 can be used to detect antibodies and antigen

CC in body fluids from individuals exposed to HCV e.g. in

CC confirmatory, competition or neutralisation assays.

XX

SQ Sequence 594 AA;

Query Match 0.5%; Score 59; DB 14; Length 594;

Best Local Similarity 34.5%; Pred. No. 4,26-62;

Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1821 VLLPAIKETVYKQIEKNNKNNHMPFMSHLQ 1849

L I I I I I I I I I I I I I I I I I I I

Db 403 viapavqtawqkietiwaklmwnflisq 431

RESULT 147

AAB69024

ID AAB69024 standard; Peptide; 594 AA.

XX

AC AAB69024;

XX

DT 17 APR 2001 (first entry)

XX

DE HCV recombinant antigen pHCV-63 amino acid sequence SEQ ID No:48.

XX Hepatitis C virus; HCV antigen; detection antibody.

XX Hepatitis C virus.

OS US6172189-H1.

XX

PN 09-JAN-2001.

XX

PD 02-JUN-1997; 97US-0867611.

XX

PR 19-NOV-1992; 92US-0989843.

XX

PR 10-JAN-1994; 94US-0179896.

XX

PR 01-MAY-1996; 96US-0646757.

XX

PR 24 APR-1990; 90US-0675832.

XX

PR 07-NOV-1990; 90US-0614069.

XX

PR 21-AUG-1991; 91US-0748561.

XX

PR 21-AUG-1991; 91US-0748566.

XX

PR 29-OCT-1991; 91US-0748565.

XX

PA (ABBO) ABBOTT LAB.

XX

XX Devare SC, Desai SM, Casey SM, Bailey SH, Dawson CL, Supprocht KP;

XX

XX WPI; 2001-122352/13.

XX

XX New recombinant antigens representing distinct antigenic regions of

XX Hepatitis C virus (HCV) genome, useful for detection of antibodies and

XX antigens in body fluids of individuals exposed to HCV.

XX

PS Claim 2; Column 177-180, 167pp; English.

XX

CC The present invention describes recombinant Hepatitis C virus (HCV) antigens (I). (i) is useful as a reagent for the detection of antibodies and antigen in body fluids from individuals exposed to HCV. The HCV assay uses reliable and efficient reagents and methods to accurately detect the presence of HCV antibodies in samples obtained from individuals suspected of having HCV infection. AAR652218 to AAR652335.

XX

CC AAB61371 to AAB61379 and AAB69001 to AAB69032 represent sequences used

XX in the exemplification of the present invention.

XX

SQ Sequence 594 AA;

Query Match 0.5%; Score 59; DB 22; Length 594;

Best Local Similarity 34.5%; Pred. No. 4,26-62;

Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1821 VLLPAIKETVYKQIEKNNKNNHMPFMSHLQ 1849

L I I I I I I I I I I I I I I I I I I I

Db 403 viapavqtawqkietiwaklmwnflisq 431

PT human melanoma cell line, tissue growth, antigen/antibody
XX and anti-inflammatory or tumor inhibition activity

XX Claim 1: Page 53-55; 206pp; English.

XX A specific antigenic region of the HIV genome is expressed as a
XX fusion protein pHV-62 can be used to detect antibodies and antigen
XX in body fluids from individuals exposed to HIV and, in
XX confirmatory, competition or neutralisation assays.

XX Sequence 592 AA:

Query Match: 0.58; Score 59; DB 14; Length 592;
Best local similarity: 45.58; Pred. Res. 4,20-024;
Matches 10; Conservative 9; Mismatches 10; Gaps 0

QY 1821 VLHVALKIKYQIEKNSWENHMCSPNSIIQ 1849
1: 111 111 1 11 11 11
R: 401 vlpavpwpqyktwakhmktlslq 429

RESULT 145

AA669023

ID AAB69023 standard; peptide: 592 AA.

AC AAB69023;

DI 17-APR-2001 (first entry)

DE HIV recombinant antigen pHV-62 amino acid sequence SEQ ID NO:47.

EE Hepatitis C virus; HIV; antigen: detection; antibody.

ES Hepatitis C virus.

FN US6172189-B1.

PD 09-JAN-2001.

PF 02-JUN-1997; 97US-0867611.

PR 19-NOV-1992; 92US-0989843.

PR 10-JAN-1994; 94US-0179896.

PR 01-MAY-1996; 96US-0646757.

PR 24-AUG-1990; 90US-0572822.

PR 07-NOV-1990; 90US-0614069.

PR 21-AUG-1991; 91US-0748561.

PR 21-AUG-1991; 91US-0748566.

PR 29-OCT-1991; 91US-0748565.

XX (ABBO) ABBOTT LAB.

XX Inverse S5, Desai SM, Casey JM, Supprent KK, Lawson GL, Galanter RA;

XX Lesniowski KK, Stewart JL, Supprent KK;

XX WPI: 2001-122352713.

XX New recombinant antigens representing distinct antigenic regions of
XX Hepatitis C virus (HCV) genome, useful for detection of antibodies and
XX antigens in body fluids of individuals exposed to HCV.

XX Example 16; Column 174-178; 167pp; English.

XX The present invention describes recombinant Hepatitis C virus (HCV)
XX antigens (1). (1) is useful as a reagent for the detection of antibodies
XX and antigen in body fluids from individuals exposed to HCV. The HCV
XX assay uses reliable and efficient reagents and methods to accurately
XX detect the presence of HCV antibodies in samples obtained from
XX individuals suspected of having HCV infection. AAB69023 to AAB69047
XX in the exemplification of the present invention.

XX Hepatitis C assay using recombinant antigenic region antigens - for
XX detection of antibodies and antigen in body fluids from individuals

Query Match: 0.58; Score 59; DB 21; Length 592;
Best local similarity: 45.58; Pred. Res. 4,20-024;
Matches 9; Conservative 10; Mismatches 16; Indels 0; Gaps 0;

QY 60 MRLHVAWQDHVITDEEEMRQRYEYHEHVALTEKLEN 605

1: 111 111 1 11 11 11 11 11

R: 306 vlpavpwpqyktwakhmktlslq 429

RESULT 144

AA669023

ID AAB69023 standard; peptide: 592 AA.

AC AAB69023;

DI 01-JUL-1996 (first entry)

DE HIV antigen, fusion protein pHV-62.

EE Hepatitis C virus; HIV; antigen: detection; antibody.

ES CMV-Koz synthesis; immunodot assay; N6; A; anti-Hepatitis.

FN Hepatitis C virus.

PD 01-JUL-1996.

PF 21-AUG-1991; 92US-0807197.

PR 21-AUG-1991; 91US-0748561.

XX (ABBO) ABBOTT LAB.

XX Casey JM, Desai SM, Inverse S5, Supprent KK;

XX WPI: 2001-122352713.

XX Hepatitis C assay using recombinant antigenic region antigens - for
XX detection of antibodies and antigen in body fluids from individuals

XX 24-JUN-1998; 98JP-0193788.
 XX (SEKK) SEIKAGAKO KOGYO CO LTD.
 XX WPI: 2000-140125/13.
 XX A hyaluronate synthase modified protein, useful as a research reagent
 PT for biochemical research and medical development
 XX Claim 12; Page: 30pp; Japanese.
 XX The present invention describes a recombinant protein which consists of
 CC three continuous regions (N-terminal region, internal region and
 CC C-terminal region) where one or two regions among the above three
 CC regions is selected from the three hyaluronate synthase (HAS) modified
 CC proteins HAS1, HAS2 and HAS3 (e.g., a protein made from HAS1 amino acids
 CC 1 to 71, HAS2 amino acids 52 to 379 and HAS3 amino acids 410 to 583).
 CC The protein is useful as a research reagent for biochemical research
 CC and medical development. The invention provides a HAS modified protein
 CC of increased or lowered activity. AAY78127 to AAY78142 represent
 CC specifically claimed examples of recombinant proteins from the present
 CC invention.
 CC N.B. The present sequence is not given in the present specification,
 CC but is derived from sequences given as specified in the claim.
 XX Sequence 553 AA:
 SQ
 Query Match 0.5%; Score 59; DP 21; Length 553;
 Best Local Similarity 28.6%; Pred No. 4002;
 Matches 14; Conservative 11; Mismatches 24; Indels 0; Gaps 0;
 QY 973 VTQQLATFEFFQPEKFKSKHCKLSETFNILSCVYSPSYIAKIMKV 1021
 DB 58 LKqLafLchukmkksldqfkkkkskclaaqyepdyfirkelpy 106
 RESULT 142
 AAY78127
 ID AAY78127 standard; Protein: 573 AA.
 XX
 AC AAY78127;
 DT 27-APR-2000 (first entry)
 XX
 DE Recombinant chimeric hyaluronate synthase modified protein #1.
 XX
 KW Mouse; murine; hyaluronate synthase; modification: HAS1; HAS2; HAS3;
 KW research reagent; biochemical research, medical development, chimeric.
 XX
 CG Chimeric, Mus sp.
 CS Synthetic.
 XX JP2930804086-A.
 XX 11-JAN-2000.
 XX 24-JUN-1998; 98JP-0193788.
 XX 24-JUN-1998; 98JP-0193788.
 XX (SEKK) SEIKAGAKO KOGYO CO LTD.
 XX WPI: 2000-140125/13.
 XX A hyaluronate synthase modified protein - useful as a research reagent
 PT for biochemical research and medical development
 XX Claim 2; Page: 30pp; Japanese.
 XX The present invention describes a recombinant protein which consists of
 CC three continuous regions (N-terminal region, internal region and

CC C-terminal region) where one or two regions among the above three
 CC regions is selected from the three hyaluronate synthase (HAS) modified
 CC proteins HAS1, HAS2 and HAS3 (e.g., a protein made from HAS1 amino acids
 CC 1 to 71, HAS2 amino acids 52 to 379 and HAS3 amino acids 410 to 583).
 CC The protein is useful as a research reagent for biochemical research
 CC and medical development. The invention provides a HAS modified protein
 CC of increased or lowered activity. AAY78127 to AAY78142 represent
 CC specifically claimed examples of recombinant proteins from the present
 CC invention.
 CC N.B. The present sequence is not given in the present specification,
 CC but is derived from sequences given as specified in the claim.
 XX Sequence 573 AA:
 SQ
 Query Match 0.5%; Score 59; DP 21; Length 573;
 Best Local Similarity 28.6%; Pred No. 4002;
 Matches 14; Conservative 11; Mismatches 24; Indels 0; Gaps 0;
 QY 973 VTQQLATFEFFQPEKFKSKHCKLSETFNILSCVYSPSYIAKIMKV 1021
 DB 78 LKqLafLchukmkksldqfkkkkskclaaqyepdyfirkelpy 106
 RESULT 143
 AAY94928
 ID AAY94928 standard; Protein: 590 AA.
 XX
 AC AAY94928;
 DT 16-JUN-2000 (first entry)
 XX
 DE Human secreted protein clone p9195.1 protein sequence SEQ ID NO:62.
 XX
 KW Human; secreted protein; immunostimulant; immunosuppressant; virucide;
 KW antitubercial; antifungal; cytostatic; antiinflammatory; dermatological;
 KW antidiabetic; antiasthmatic; antiarthritic; antineumatic; protozoacide;
 KW antithyroid; immune deficiency; severe combined immunodeficiency; SCID;
 KW infection, HIV, hepatitis, malaria; autoimmune disorder; systemic lupus;
 KW connective tissue disease; multiple sclerosis; erythematosis;
 KW rheumatoid arthritis; autoimmune pulmonary inflammation; asthma;
 KW Guillain-Barre syndrome; autoimmune thyroiditis; myasthenia gravis;
 KW insulin dependent diabetes mellitus; graft-versus-host-disease;
 KW autoimmune inflammatory eye disease; allergy.
 CG Homo sapiens.
 XX WO200009552-A1.
 XX 24-FEB-2000.
 XX 13-AUG-1999; 99WO-018298.
 XX 14-AUG-1998; 98US-0096622.
 XX 17-AUG-1998; 98US-0096815.
 XX 04-SEP-1998; 98US-0099229.
 XX 23-OCT-1998; 98US-0105368.
 XX 08-JAN-1999; 98US-0115234.
 XX 12-FEB-1999; 98US-0119431.
 XX 18-FEB-1999; 98US-0120575.
 XX 30-APR-1999; 98US-0132020.
 XX 11-AUG-1999; 98US-0096622.
 XX (GEMV) GENETICS INST INC.
 XX Jacobs K, McCoy JM, LaVallic ER, Collins-Paché JA, Evans G;
 PI Merberg D, Treacy M, Aostino MJ, Steininger RJ, Spaulding V;
 PI Wong GG, Clark HF, Fechtel K;
 XX WPI: 2000-205079/18.
 XX Few polynucleotides encoding secreted proteins, which may have e.g.,
 PT nutritional, chemokine, immune stimulating or suppressing,


```

XX New A non B hepatitis, probe, vaccine, diagnosis;
KW passive immunotherapy; antigen.
XX Hepatitis C virus.
XX Key Location/Qualifiers
FH Region 1..27
FT /label: 376
FT Region 28..42
FT /label: adaptor
FT Region 43..377
FT /label: NAB
FT Region 378..342
FT /label: extra
XX EP112216 A.
XX 31 MAY 1989.
XX 18-NOV-1988; 88EP-0310922.
XX 14-NOV 1988; 88US 0271450.
XX (CHIR-) CHIRON CORP.
XX Houghton M, Choo QL, Kuo G;
XX WP1: 1989-159274/22.
XX N-PSDB; AAN92104.
XX Purified hepatitis C virus -
XX and assorted nucleic acids and polypeptide(s)
XX Example, Fig 46.1 % 2. 199pp, English.
XX Purified hepatitis C virus (HCV) and purified or recombinant HCV nucleic
XX acids (NAs), encoding HCV polypeptides or epitopes, and polypeptides
XX are claimed. HCV is a causative agent of non A, non B hepatitis (NANBH).
XX The NAs may be used to design probes for detection of HCV NAs in samples.
XX The polypeptides may be used as immun assay reagents and vaccines, and
XX to produce antibodies useful for diagnosis and passive immunotherapy.
XX The purified virus may also be used in vaccines.
XX Sequence 382 AA;
XX
XX Query Match 9.5%; Score 59; DB 10; Length 382;
XX Post Local Similarity 34.5%; Pref. No. 2,200,027;
XX Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;
XX
XX 182) VITADREYPCIIHMKKMPGDSKSLQ DTA
XX 191 VTAQQTWQKLTWKNMFIISLQ 219
XX
XX RESULT 124
XX AAG15375
XX ID AAG15375 standard; protein; 414 AA.
XX AC AAG15375;
XX DF 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 15604.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX Eukaryotic translation; gene expression; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX EP1034405-A2.
XX
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0133180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0135788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 14-APR-1999; 99US-0129845.
XX 21-APR-1999; 99US-0130077.
XX 23-APR-1999; 99US-0130449.
XX 28-APR-1999; 99US-0130510.
XX 30-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136292.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137272.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 24-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140645.
XX 24-JUN-1999; 99US-0140823.
XX 28-JUN-1999; 99US-0140991.
XX 30-JUN-1999; 99US-0141287.
XX 01-JUL-1999; 99US-0141842.
XX 01-JUL-1999; 99US-0142154.

```


PR 19-APR-1999; 9905-0140077; PR 19-JUL-1999; 9905-0144335.
PR 21-APR-1999; 9905-0140449; PR 20-JUL-1999; 9905-0144352.
PR 24-APR-1999; 9905-0140510; PR 20-JUL-1999; 9905-0144632.
PR 24-APR-1999; 9905-0140893; PR 20-JUL-1999; 9905-0144884.
PR 28-APR-1999; 9905-0141449; PR 21-JUL-1999; 9905-0144814.
PR 30-APR-1999; 9905-0142048; PR 21-JUL-1999; 9905-0145086.
PR 30-APR-1999; 9905-0142407; PR 21-JUL-1999; 9905-0145088.
PR 04-MAY-1999; 9905-0142424; PR 22-JUL-1999; 9905-0145087.
PR 05-MAY-1999; 9905-0142485; PR 22-JUL-1999; 9905-0145089.
PR 06-MAY-1999; 9905-0142486; PR 22-JUL-1999; 9905-0145192.
PR 06-MAY-1999; 9905-0142487; PR 23-JUL-1999; 9905-0145145.
PR 07-MAY-1999; 9905-0142863; PR 23-JUL-1999; 9905-0145218.
PR 11-MAY-1999; 9905-0142556; PR 23-JUL-1999; 9905-0145224.
PR 14-MAY-1999; 9905-0142518; PR 26-JUL-1999; 9905-0145276.
PR 14-MAY-1999; 9905-0142519; PR 27-JUL-1999; 9905-0145913.
PR 14-MAY-1999; 9905-0144221; PR 27-JUL-1999; 9905-0145918.
PR 14-MAY-1999; 9905-0144370; PR 27-JUL-1999; 9905-0145919.
PR 18-MAY-1999; 9905-0144768; PR 28-JUL-1999; 9905-0145951.
PR 18-MAY-1999; 9905-0144941; PR 02-AUG-1999; 9905-0146386.
PR 20-MAY-1999; 9905-0145123; PR 02-AUG-1999; 9905-0146388.
PR 21-MAY-1999; 9905-0145353; PR 02-AUG-1999; 9905-0146389.
PR 24-MAY-1999; 9905-0145532; PR 03-AUG-1999; 9905-0147038.
PR 25-MAY-1999; 9905-0146021; PR 04-AUG-1999; 9905-0147204.
PR 27-MAY-1999; 9905-0146492; PR 04-AUG-1999; 9905-0147302.
PR 28-MAY-1999; 9905-0146782; PR 05-AUG-1999; 9905-0147192.
PR 01-JUN-1999; 9905-0147222; PR 05-AUG-1999; 9905-0147260.
PR 03-JUN-1999; 9905-0147528; PR 06-AUG-1999; 9905-0147303.
PR 04-JUN-1999; 9905-0147502; PR 06-AUG-1999; 9905-0147416.
PR 07-JUN-1999; 9905-0147723; PR 06-AUG-1999; 9905-0147493.
PR 08-JUN-1999; 9905-0148094; PR 09-AUG-1999; 9905-0147935.
PR 10-JUN-1999; 9905-0148540; PR 10-AUG-1999; 9905-0148171.
PR 10-JUN-1999; 9905-0148847; PR 11-AUG-1999; 9905-0148319.
PR 14-JUN-1999; 9905-0149119; PR 12-AUG-1999; 9905-0148341.
PR 16-JUN-1999; 9905-0149452; PR 13-AUG-1999; 9905-0148565.
PR 16-JUN-1999; 9905-0149453; PR 13-AUG-1999; 9905-0148684.
PR 17-JUN-1999; 9905-0149492; PR 16-AUG-1999; 9905-0149368.
PR 18-JUN-1999; 9905-0149454; PR 17-AUG-1999; 9905-0149175.
PR 18-JUN-1999; 9905-0149455; PR 18-AUG-1999; 9905-0149426.
PR 18-JUN-1999; 9905-0149456; PR 20-AUG-1999; 9905-0149722.
PR 18-JUN-1999; 9905-0149457; PR 20-AUG-1999; 9905-0149723.
PR 18-JUN-1999; 9905-0149458; PR 20-AUG-1999; 9905-0149929.
PR 18-JUN-1999; 9905-0149459; PR 23-AUG-1999; 9905-0149902.
PR 18-JUN-1999; 9905-0149460; PR 23-AUG-1999; 9905-0149930.
PR 18-JUN-1999; 9905-0149461; PR 25-AUG-1999; 9905-0150566.
PR 18-JUN-1999; 9905-0149462; PR 26-AUG-1999; 9905-0150884.
PR 18-JUN-1999; 9905-0149463; PR 27-AUG-1999; 9905-0151065.
PR 18-JUN-1999; 9905-0149464; PR 27-AUG-1999; 9905-0151066.
PR 18-JUN-1999; 9905-0149465; PR 27-AUG-1999; 9905-0151080.
PR 18-JUN-1999; 9905-0149466; PR 30-AUG-1999; 9905-0151303.
PR 18-JUN-1999; 9905-0149467; PR 31-AUG-1999; 9905-0151438.
PR 18-JUN-1999; 9905-0149468; PR 01-SEP-1999; 9905-0151930.
PR 18-JUN-1999; 9905-0149469; PR 07-SEP-1999; 9905-0152363.
PR 18-JUN-1999; 9905-0149470; PR 10-SEP-1999; 9905-0153070.
PR 18-JUN-1999; 9905-0149471; PR 13-SEP-1999; 9905-0153758.
PR 18-JUN-1999; 9905-0149472; PR 15-SEP-1999; 9905-0154018.
PR 18-JUN-1999; 9905-0149473; PR 16-SEP-1999; 9905-0154039.
PR 18-JUN-1999; 9905-0149474; PR 20-SEP-1999; 9905-0154779.
PR 18-JUN-1999; 9905-0149475; PR 22-SEP-1999; 9905-0155139.
PR 18-JUN-1999; 9905-0149476; PR 23-SEP-1999; 9905-0155486.
PR 18-JUN-1999; 9905-0149477; PR 24-SEP-1999; 9905-0155659.
PR 18-JUN-1999; 9905-0149478; PR 28-SEP-1999; 9905-0156458.
PR 18-JUN-1999; 9905-0149479; PR 29-SEP-1999; 9905-0156896.
PR 18-JUN-1999; 9905-0149480; PR 04-OCT-1999; 9905-0157117.
PR 18-JUN-1999; 9905-0149481; PR 05-OCT-1999; 9905-0157753.
PR 18-JUN-1999; 9905-0149482; PR 06-OCT-1999; 9905-0157865.
PR 18-JUN-1999; 9905-0149483; PR 07-OCT-1999; 9905-0158029.
PR 18-JUN-1999; 9905-0149484; PR 08-OCT-1999; 9905-0158232.
PR 18-JUN-1999; 9905-0149485; PR 12-OCT-1999; 9905-0158369.
PR 18-JUN-1999; 9905-0149486; PR 13-OCT-1999; 9905-0159293.
PR 18-JUN-1999; 9905-0149487; PR 13-OCT-1999; 9905-0159294.
PR 18-JUN-1999; 9905-0149488; PR 13-OCT-1999; 9905-0159295.
PR 18-JUN-1999; 9905-0149489; PR 14-OCT-1999; 9905-0159329.

AA	(GETH) GENETECH INC.	PE	23-MAR-1999;
XX	Botstein D, Finkblatt T,	PE	99US-0125264
XX	WPI: 1992-167169,266.	PE	99US-0126785
XX	N-PSUB: AN24143.	PE	99US-0127462
XX		PE	99US-0128234
XX		PE	99US-0128714
XX		PE	99US-0129445
XX		PE	99US-0130077
XX		PE	99US-0130449
XX		PE	99US-0130510
XX		PE	99US-0130891
XX		PE	99US-0131449
XX		PE	99US-0132048
XX		PE	99US-0132407
XX		PE	99US-0132484
XX		PE	99US-0132485
XX		PE	99US-0132486
XX		PE	99US-0132487
XX		PE	99US-0132863
XX		PE	99US-0134256
XX		PE	99US-0134318
XX		PE	99US-0134319
XX		PE	99US-0134321
XX		PE	99US-0134370
XX		PE	99US-0134376
XX		PE	99US-0134941
XX		PE	99US-0135124
XX		PE	99US-0135353
XX		PE	99US-0135629
XX		PE	99US-0136021
XX		PE	99US-0136392
XX		PE	99US-0136782
XX		PE	99US-0137222
XX		PE	99US-0137528
XX		PE	99US-0137502
XX		PE	99US-0137724
XX		PE	99US-0138094
XX		PE	99US-0138540
XX		PE	99US-0138847
XX		PE	99US-0139119
XX		PE	99US-0139452
XX		PE	99US-0139453
XX		PE	99US-0139492
XX		PE	99US-0139454
XX		PE	99US-0139455
XX		PE	99US-0139456
XX		PE	99US-0139457
XX		PE	99US-0139458
XX		PE	99US-0139459
XX		PE	99US-0139460
XX		PE	99US-0139461
XX		PE	99US-0139462
XX		PE	99US-0139463
XX		PE	99US-0139750
XX		PE	99US-0139763
XX		PE	99US-0139817
XX		PE	99US-0139849
XX		PE	99US-0140354
XX		PE	99US-0140695
XX		PE	99US-0140823
XX		PE	99US-0140991
XX		PE	99US-0141287
XX		PE	99US-0141842
XX		PE	99US-0142154
XX		PE	99US-0142055
XX		PE	99US-0142390
XX		PE	99US-0142803
XX		PE	99US-0142920
XX		PE	99US-0142977
XX		PE	99US-0143542
XX		PE	99US-0144005
XX		PE	99US-0144025

Query Match 0.5%; Score 59; Db 13; Length 246;
 Best Local Similarity 22.7%; Pred. No. 1,56,02;
 Matches 15; Conservative 15; Mismatches 36; Indels 0; Gaps 0;
 1772 GPEPSPGALHLLSQVRLLEPTEAGSASGALSLSEKELATLAPVLAAPKPYV 1831
 143 pGeflatthumadhrilaprepcisipgdrptimpmamtlirklltpatalar 202
 1842 QLEKRN 1837
 203 qllrdw 209
 RESULT 116
 AAG20941
 ID AAG20941 standard; Protein: 289 AA.
 AC AAG20941;
 XX 17-OCT-2000 (first entry)
 XX Arabidopsis thaliana protein fragment SEQ ID NO. 24202.
 DE Arabidopsis thaliana.
 PS Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX Arabidopsis thaliana.
 PS Arabidopsis thaliana.
 FN EP103405-A2.
 IN 06-SEP-2000.
 XX 25-FEB-2000, 2600PR 930439.
 XX 25-FEB-1999; 99US-0121925.
 XX 05-MAR-1999; 99US-0121930.
 XX 09-MAR-1999; 99US-0124548.

PI Translating, useful for expressed gene products and for controlling

XX expressed, of a target gene.

XX Team 14: Page 540: 23pp: English.

XX The present invention describes ribes polynucleotides, such as complete cDNA

XX sequences and/or sequences of genomic DNA encompassing complete genes,

XX portions of genes, and/or intergenic regions, collectively referred to

XX as sequence determined DNA fragments (SDFs), from corn plants and

XX Arabidopsis thaliana. The SDFs are promoters, structural genes, and

XX untranslated regions (UTRs), or 5' to 3' terminal sequences. They can be

XX used for expressing a gene product, and controlling expression of a

XX target gene, either as a promoter, a structural gene, an HIR or as a

XX c-terminus sequence. They are also useful as tools for genetic

XX mapping, and identification of a particular individual plant or for

XX clustering a group of plants with a common trait. AAA78433 to AAA78640

XX are Arabidopsis cDNA sequences and polypeptides encoded by them. In the

XX present invention, they represent the specifically claimed

XX sequence. 234 AA.

Query Match 0.58; Score 59; Db 21; Length 234;

Best Local Similarity 43.88; Prod. No. 110-02;

Matches 14; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

27 433 LRGISSTYVSSLEFRLMLHUVS11HVV13 54

11 | : | | | | | | | : | : | | | | |

16 203 LPPITLDSPLATAVTIGSIVHASTQ 232

11 | : | | | | | | | : | : | | | | |

RESULT 110

AAAP90138

ID AAP90138 standard; protein; 237 AA.

XX AAAP90138;

XX 21 NOV-1989 (first entry)

XX Peptide encoded by combined cDNA of hepatitis C virus cDNAs in

XX clones 36 and 81.

XX Hepatitis C virus; open reading frame; clone 36; clone 81; protein

XX vaccine.

XX Pan troglodytes.

XX GH212511-A.

XX 26-JUL-1989.

XX 18-NOV-1988; BBGB-0027024.

XX 18-NOV-1987; H7US-0122714.

XX (CHIR) CHIRON CORPORATION.

XX Hershman M, Choo QL, Kuo G;

XX WJ; 1989; 215054/40.

XX N-DSUB; AAN90407.

XX Hepatitis C virus gene; used for prodn. of polynucleotide probes,

XX polypeptide(s) and antibodies for diagnosis, prevention and treatment

XX of infection.

XX Disclosure; fig 6; 235pp; English.

XX The sequence is the peptide encoded by the combined open reading frame

XX of hepatitis C virus (HCV) cDNAs in clones 36 and 81 (see AAN90407).

XX The polypeptides are used to diagnose HCV induced NANBH, to raise

XX antibodies for immunoassay or treatment, or to produce vaccines.

XX Sequence 237 AA;

XX Query Match 0.58; Score 59; Db 21; Length 237;

XX Best Local Similarity 24.58; Prod. No. 110-02;

XX Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

XX 1921 VLLIAKPTVPCITPYSWQHREGRPSILQ 1849

XX The present invention describes polynucleotides, such as complete cDNA

XX sequences and/or sequences of genomic DNA encompassing complete genes,

XX portions of genes, and/or intergenic regions, collectively referred to

XX as sequence determined DNA fragments (SDFs), from corn plants and

XX Arabidopsis thaliana. The SDFs are promoters, structural genes, and

XX untranslated regions (UTRs), or 5' to 3' terminal sequences. They can be

XX used for expressing a gene product, and controlling expression of a

XX target gene, either as a promoter, a structural gene, an HIR or as a

XX c-terminus sequence. They are also useful as tools for genetic

PR 16-JUL-1999; 990S-0144085.
 PR 16-JUL-1999; 990S-0144086.
 PR 19-JUL-1999; 990S-0144325.
 PR 19-JUL-1999; 990S-0144331.
 PR 19-JUL-1999; 990S-0144332.
 PR 19-JUL-1999; 990S-0144333.
 PR 19-JUL-1999; 990S-0144334.
 PR 19-JUL-1999; 990S-0144335.
 PR 20-JUL-1999; 990S-0144352.
 PR 20-JUL-1999; 990S-0144362.
 PR 20-JUL-1999; 990S-0144864.
 PR 21-JUL-1999; 990S-0144814.
 PR 21-JUL-1999; 990S-0145086.
 PR 21-JUL-1999; 990S-0145089.
 PR 22-JUL-1999; 990S-0145085.
 PR 22-JUL-1999; 990S-0145087.
 PR 22-JUL-1999; 990S-0145089.
 PR 22-JUL-1999; 990S-0145192.
 PR 23-JUL-1999; 990S-0145145.
 PR 23-JUL-1999; 990S-0145218.
 PR 23-JUL-1999; 990S-0145224.
 PR 26-JUL-1999; 990S-0145276.
 PR 27-JUL-1999; 990S-0145913.
 PR 27-JUL-1999; 990S-0145918.
 PR 27-JUL-1999; 990S-0145919.
 PR 28-JUL-1999; 990S-0145951.
 PR 02-AUG-1999; 990S-0146386.
 PR 02-AUG-1999; 990S-0146388.
 PR 02-AUG-1999; 990S-0146389.
 PR 03-AUG-1999; 990S-0147038.
 PR 04-AUG-1999; 990S-0147204.
 PR 04-AUG-1999; 990S-0147302.
 PR 05-AUG-1999; 990S-0147192.
 PR 05-AUG-1999; 990S-0147260.
 PR 06-AUG-1999; 990S-0147303.
 PR 06-AUG-1999; 990S-0147304.
 PR 09-AUG-1999; 990S-0147393.
 PR 09-AUG-1999; 990S-0147395.
 PR 10-AUG-1999; 990S-0148171.
 PR 11-AUG-1999; 990S-0148319.
 PR 12-AUG-1999; 990S-0148341.
 PR 13-AUG-1999; 990S-0148565.
 PR 13-AUG-1999; 990S-0148684.
 PR 16-AUG-1999; 990S-0149368.
 PR 17-AUG-1999; 990S-0149175.
 PR 18-AUG-1999; 990S-0149426.
 PR 20-AUG-1999; 990S-0149722.
 PR 20-AUG-1999; 990S-0150721.
 PR 20-AUG-1999; 990S-0149929.
 PR 23-AUG-1999; 990S-0149902.
 PR 23-AUG-1999; 990S-0149930.
 PR 25-AUG-1999; 990S-0150566.
 PR 26-AUG-1999; 990S-0150884.
 PR 27-AUG-1999; 990S-0151065.
 PR 27-AUG-1999; 990S-0151066.
 PR 27-AUG-1999; 990S-0151080.
 PR 30-AUG-1999; 990S-0151303.
 PR 31-AUG-1999; 990S-0151338.
 PR 01-SEP-1999; 990S-0151930.
 PR 07-SEP-1999; 990S-0152363.
 PR 10-SEP-1999; 990S-0153070.
 PR 13-SEP-1999; 990S-0153758.
 PR 15-SEP-1999; 990S-0154018.
 PR 16-SEP-1999; 990S-0154039.
 PR 20-SEP-1999; 990S-0154779.
 PR 22-SEP-1999; 990S-0155139.
 PR 23-SEP-1999; 990S-0155486.
 PR 24-SEP-1999; 990S-0155659.
 PR 28-SEP-1999; 990S-0156458.
 PR 29-SEP-1999; 990S-0156596.
 PR 04-OCT-1999; 990S-0157117.
 PR 05-OCT-1999; 990S-0157753.
 PR 06-OCT-1999; 990S-0157865.

PR 07-OCT-1999; 990S-0158029.
 PR 08-OCT-1999; 990S-0158232.
 PR 12-OCT-1999; 990S-0158369.
 PR 13-OCT-1999; 990S-0159293.
 PR 13-OCT-1999; 990S-0159294.
 PR 13-OCT-1999; 990S-0159295.
 PR 14-OCT-1999; 990S-0159329.
 PR 14-OCT-1999; 990S-0159330.
 PR 14-OCT-1999; 990S-0159331.
 PR 14-OCT-1999; 990S-0159637.
 PR 14-OCT-1999; 990S-0159638.
 PR 18-OCT-1999; 990S-0159584.
 PR 21-OCT-1999; 990S-0160741.
 PR 21-OCT-1999; 990S-0160767.
 PR 21-OCT-1999; 990S-0160768.
 PR 21-OCT-1999; 990S-0160770.
 PR 21-OCT-1999; 990S-0160814.
 PR 21-OCT-1999; 990S-0160815.
 PR 22-OCT-1999; 990S-0160980.
 PR 22-OCT-1999; 990S-0160981.
 PR 22-OCT-1999; 990S-0160989.
 PR 25-OCT-1999; 990S-0161404.
 PR 25-OCT-1999; 990S-0161405.
 PR 25-OCT-1999; 990S-0161406.
 PR 26-OCT-1999; 990S-0161359.
 PR 26-OCT-1999; 990S-0161360.
 PR 26-OCT-1999; 990S-0161361.
 PR 26-OCT-1999; 990S-0161920.
 PR 28-OCT-1999; 990S-0161992.
 PR 28-OCT-1999; 990S-0161993.
 PR 29-OCT-1999; 990S-0162142.

Query Match 0.58; Score 59; DB 21; Length 220;
 Best Local Similarity 35.7%; Pred. No. 99;
 Matches 15; Conservative 9; Mismatches 18; Indels 0; Gaps 0;

QY 736 KKKKLSVITAVIPSEWHLEHIFKGLIVGLAHVYVSHNE 777
 EL 139 KKKKLSVITAVIPSEWHLEHIFKGLIVGLAHVYVSHNE 777

RESULT 108
 AAB24905
 ID AAB24905 standard; Peptide: 232 AA.
 XX AAB24905;
 AC AAB24905;
 XX AAB24905;
 PT 27-NOV-2000 (first entry)
 XX
 DE Plant SDF encoded polypeptide sequence SDF list 1 30:421.
 XX
 KW Plant, corn, Arabidopsis thaliana, sequence determined DNA fragment;
 KW SDF; genetic mapping; identification; promoter; structural gene; UTR;
 KW untranslated region, expression control.
 XX
 OS plant.
 XX
 PN W0269040695-A2.
 XX
 PD 13-JUL-2000.
 XX
 PF 07-JAN-2000; 2000W0-US00466.
 XX
 PR 08-JAN-1999; 990S-0115293.
 XX
 PA (CERE-) CERES INC.
 XX
 PT Alexander N. Proyer V. Chen X. Subramanian G. Troukhan MG;
 PT Zheng I;
 XX
 PP WPI; 2000-465970/40.
 XX
 PT New corn plant and Arabidopsis thaliana sequence-determined DNA

Query Match 0.5%; Score 59; DB 10; Length 117;
 Best Local Similarity 34.5%; Pred. No. 40;
 Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;
 QY 1821 VLLPALKTKYKQIEKNWKNHMFPMSTIQ 1849
 I: III: III: I II I: I I
 Db 57 viapavqtwnqkietfwakmwnfisiq 85
 RESULT 101
 AAR13354
 ID AAR13354 standard; protein; 117 AA.
 XX AC AAR13354;
 XX DT 02-MAR-1990 (first entry)
 XX DE Sequence of the polypeptide encoded in the hepatitis C virus (HCV) cDNA
 XX DE insert in clone 81.
 XX KW Hepatitis C virus (HCV); non-A, non-B hepatitis (HANBB)
 XX OS Hepatitis C virus.
 XX FN EP318216-A.
 XX PD 31-MAY-1989.
 XX PF 18-NOV-1988; 88EP-0310922.
 XX PR 14-NOV-1988; 88US-0271450; US-122714.
 XX PA (CHIR) CHIRON CORP.
 XX PI Boughton M, Choo Q L, Kao G,
 XX WPI; 1989-159274/22.
 XX R-1338; AAR13354.
 XX PT Purified hepatitis C virus
 XX PT - and associated nucleic acids and polypeptide(s)
 XX PS Claim 13; Figure 4; 13pp; English.
 XX CC It is the putative sequence encoded in the open reading frame of
 XX CC hepatitis C virus (HCV) cDNA insert in clone 81. It is an epitope,
 XX CC portions of which could be used as immunoassay reagents and vaccines and
 XX CC to generate antibodies useful in diagnosis and passive immunotherapy for
 XX CC HCV infection/non-A, non-B hepatitis.
 XX SQ Sequence 117 AA;
 Query Match 0.5%; Score 59; DB 10; Length 117;
 Best Local Similarity 34.5%; Pred. No. 40;
 Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;
 QY 1821 VLLPALKTKYKQIEKNWKNHMFPMSTIQ 1849
 I: III: III: I II I: I I
 Db 57 viapavqtwnqkietfwakmwnfisiq 85
 RESULT 102
 AAR133872
 ID AAR133872 standard; peptide; 117 AA.
 XX AC AAR133872;
 XX DT 19-JUN-1993 (first entry)
 XX DE Polypeptide p1689 comprising HCV viral antigen.
 XX

KW Hepatitis C virus; NANBH; assay; antibody; p380-JH1; p380-J; p380-J3;
 KW p408.
 XX Synthetic.
 XX FN W09206247-A.
 XX PD 01-APR-1993.
 XX PF 16-SEP-1992; 92WO-US07813.
 XX PR 16-SEP-1991; 91US-0760292
 XX PA (ABBO) ABBOTT LAB.
 XX PI Lesniowski RR, Leung TK;
 XX WPI; 1993-117563/14.
 XX DE Assay for detecting presence of antibody to hepatitis C viral
 XX DE antigen - by contacting sample with polypeptide contg. at least
 XX DE one epitope of virus antigen
 XX PS Disclosure; Page 13; 63pp; English.
 XX CC The synthetic peptide p1689 represents amino acid residues 1689-1805 of
 XX CC the hepatitis C viral antigen. The peptide may be used in an assay to
 XX CC detect antibodies to HCV and thus to diagnose chronic HCV infection.
 XX CC See also AAR133861-87.
 XX SQ Sequence 117 AA;
 Query Match 0.5%; Score 59; DB 14; Length 117;
 Best Local Similarity 34.5%; Pred. No. 40;
 Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;
 QY 1821 VLLPALKTKYKQIEKNWKNHMFPMSTIQ 1849
 I: III: III: I II I: I I
 Db 57 viapavqtwnqkietfwakmwnfisiq 85
 RESULT 103
 AAR13354
 ID AAR13354 standard; Protein; 118 AA.
 XX AC AAR13354;
 XX DT 23 OCT-1991 (first entry)
 XX DE p1689 HCV antigen (1689-1805).
 XX KW C100-3; hepatitis C virus; immunoassay; epitope.
 XX OS Synthetic.
 XX FN AP9068300-A.
 XX PD 27 JUN-1991.
 XX PF 21-DEC-1990; 90AU-0068390.
 XX PR 07-NOV-1990; 90US-0610180.
 XX PR 22-DEC-1989; 89US-0456152.
 XX PA (ABBO) ABBOTT LABORATORIES.
 XX WPI; 1991-238393/33.
 XX DE Immunological assays for hepatitis C virus antibody - by using
 XX DE polypeptide(s) contg. epitope(s) of hepatitis C virus antigens
 XX Claim 1; Page 45; 62pp; English.


```

PR 25 FEB 1999 9908-0142920.
XX PR 12-JUL-1999 9908-0142977.
PR 13-JUL-1999 9908-0143542.
PR 14-JUL-1999 9908-0143624.
PR 15-JUL-1999 9908-0144005.
PR 16-JUL-1999 9908-0144085.
PR 17-JUL-1999 9908-0144086.
PR 18-JUL-1999 9908-0144425.
PR 19-JUL-1999 9908-0144441.
PR 19-JUL-1999 9908-0144442.
PR 19-JUL-1999 9908-0144443.
PR 19-JUL-1999 9908-0144444.
PR 19-JUL-1999 9908-0144445.
PR 20-JUL-1999 9908-0144452.
PR 20-JUL-1999 9908-0144632.
PR 20-JUL-1999 9908-0144884.
PR 21-JUL-1999 9908-0144814.
PR 21-JUL-1999 9908-0145086.
PR 21-JUL-1999 9908-0145088.
PR 22-JUL-1999 9908-0145085.
PR 22-JUL-1999 9908-0145087.
PR 22-JUL-1999 9908-0145089.
PR 22-JUL-1999 9908-0145192.
PR 23-JUL-1999 9908-0145145.
PR 23-JUL-1999 9908-0145218.
PR 23-JUL-1999 9908-0145224.
PR 26-JUL-1999 9908-0145276.
PR 27-JUL-1999 9908-0145613.
PR 27-JUL-1999 9908-0145918.
PR 27-JUL-1999 9908-0145919.
PR 28-JUL-1999 9908-0145951.
PR 02-AUG-1999 9908-0146386.
PR 02-AUG-1999 9908-0146388.
PR 02-AUG-1999 9908-0146389.
PR 04-AUG-1999 9908-0147038.
PR 04-AUG-1999 9908-0147204.
PR 04-AUG-1999 9908-0147302.
PR 05-AUG-1999 9908-0147192.
PR 05-AUG-1999 9908-0147260.
PR 06-AUG-1999 9908-0147403.
PR 06-AUG-1999 9908-0147416.
PR 06-AUG-1999 9908-0147433.
PR 09-AUG-1999 9908-0147935.
PR 10-AUG-1999 9908-0148171.
PR 11-AUG-1999 9908-0148419.
PR 12-AUG-1999 9908-0148441.
PR 14-AUG-1999 9908-0148565.
PR 14-AUG-1999 9908-0148684.
PR 16-AUG-1999 9908-0149368.
PR 17-AUG-1999 9908-0149175.
PR 18-AUG-1999 9908-0149426.
PR 20-AUG-1999 9908-0149722.
PR 23-AUG-1999 9908-0149723.
PR 20-AUG-1999 9908-0149929.
PR 24-AUG-1999 9908-0149902.
PR 24-AUG-1999 9908-0149930.
PR 25-AUG-1999 9908-0150566.
PR 26-AUG-1999 9908-0150884.
PR 27-AUG-1999 9908-0151065.
PR 27-AUG-1999 9908-0151066.
PR 27-AUG-1999 9908-0151080.
PR 30-AUG-1999 9908-0151303.
PR 31-AUG-1999 9908-0151348.
PR 01-SEP-1999 9908-0151930.
PR 07-SEP-1999 9908-0152363.
PR 10-SEP-1999 9908-0153070.
PR 14-SEP-1999 9908-0153758.
PR 15-SEP-1999 9908-0154018.
PR 16-SEP-1999 9908-0154039.
PR 20-SEP-1999 9908-0154779.
PR 22-SEP-1999 9908-0155139.
PR 24-SEP-1999 9908-0155486.
PR 24-SEP-1999 9908-0155659.

```


PR 07-JUN-1999; 990S-0137724.
 PR 08-JUN-1999; 990S-0138094.
 PR 10-JUN-1999; 990S-0138540.
 PR 10-JUN-1999; 990S-0138647.
 PR 14-JUN-1999; 990S-0139119.
 PR 16-JUN-1999; 990S-0139452.
 PR 16-JUN-1999; 990S-0139453.
 PR 17-JUN-1999; 990S-0139492.
 PR 18-JUN-1999; 990S-0139454.
 PR 18-JUN-1999; 990S-0139455.
 PR 18-JUN-1999; 990S-0139456.
 PR 18-JUN-1999; 990S-0139457.
 PR 18-JUN-1999; 990S-0139458.
 PR 18-JUN-1999; 990S-0139459.
 PR 18-JUN-1999; 990S-0139460.
 PR 18-JUN-1999; 990S-0139461.
 PR 18-JUN-1999; 990S-0139462.
 PR 18-JUN-1999; 990S-0139463.
 PR 18-JUN-1999; 990S-0139464.
 PR 18-JUN-1999; 990S-0139750.
 PR 18-JUN-1999; 990S-0139763.
 PR 21-JUN-1999; 990S-0139817.
 PR 22-JUN-1999; 990S-0139869.
 PR 23-JUN-1999; 990S-0140354.
 PR 23-JUN-1999; 990S-0140354.
 PR 24-JUN-1999; 990S-0140354.
 PR 28-JUN-1999; 990S-0140824.
 PR 29-JUN-1999; 990S-0140991.
 PR 30-JUN-1999; 990S-0141287.
 PR 01-JUL-1999; 990S-0141842.
 PR 01-JUL-1999; 990S-0142154.
 PR 02-JUL-1999; 990S-0142055.
 PR 06-JUL-1999; 990S-0142390.
 PR 08-JUL-1999; 990S-0142804.
 PR 09-JUL-1999; 990S-0142929.
 PR 12-JUL-1999; 990S-0142977.
 PR 13-JUL-1999; 990S-0143542.
 PR 14-JUL-1999; 990S-0143624.
 PR 15-JUL-1999; 990S-0144005.
 PR 16-JUL-1999; 990S-0144085.
 PR 16-JUL-1999; 990S-0144086.
 PR 19-JUL-1999; 990S-0144325.
 PR 19-JUL-1999; 990S-0144331.
 PR 19-JUL-1999; 990S-0144332.
 PR 19-JUL-1999; 990S-0144333.
 PR 19-JUL-1999; 990S-0144334.
 PR 19-JUL-1999; 990S-0144335.
 PR 19-JUL-1999; 990S-0144352.
 PR 20-JUL-1999; 990S-0144632.
 PR 20-JUL-1999; 990S-0144884.
 PR 21-JUL-1999; 990S-0144814.
 PR 21-JUL-1999; 990S-0145086.
 PR 21-JUL-1999; 990S-0145088.
 PR 22-JUL-1999; 990S-0145085.
 PR 22-JUL-1999; 990S-0145087.
 PR 22-JUL-1999; 990S-0145089.
 PR 22-JUL-1999; 990S-0145192.
 PR 23-JUL-1999; 990S-0145145.
 PR 23-JUL-1999; 990S-0145218.
 PR 23-JUL-1999; 990S-0145224.
 PR 26-JUL-1999; 990S-0145276.
 PR 27-JUL-1999; 990S-0145513.
 PR 27-JUL-1999; 990S-0145918.
 PR 27-JUL-1999; 990S-0145919.
 PR 28-JUL-1999; 990S-0145951.
 PR 02-AUG-1999; 990S-0146386.
 PR 02-AUG-1999; 990S-0146388.
 PR 02-AUG-1999; 990S-0146389.
 PR 03-AUG-1999; 990S-0147038.
 PR 04-AUG-1999; 990S-0147264.
 PR 05-AUG-1999; 990S-0147802.
 PR 05-AUG-1999; 990S-0147192.
 PR 05-AUG-1999; 990S-0147260.
 PR 05-AUG-1999; 990S-0147303.

PR 06-AUG-1999; 990S-0147416.
 PR 09-AUG-1999; 990S-0147493.
 PR 09-AUG-1999; 990S-0147935.
 PR 10-AUG-1999; 990S-0148171.
 PR 11-AUG-1999; 990S-0148319.
 PR 12-AUG-1999; 990S-0148341.
 PR 13-AUG-1999; 990S-0148565.
 PR 13-AUG-1999; 990S-0148684.
 PR 16-AUG-1999; 990S-0149368.
 PR 17-AUG-1999; 990S-0149175.
 PR 18-AUG-1999; 990S-0149426.
 PR 20-AUG-1999; 990S-0149722.
 PR 20-AUG-1999; 990S-0149723.
 PR 20-AUG-1999; 990S-0149929.
 PR 23-AUG-1999; 990S-0149940.
 PR 23-AUG-1999; 990S-0149940.
 PR 25-AUG-1999; 990S-0150566.
 PR 26-AUG-1999; 990S-0150884.
 PR 27-AUG-1999; 990S-0151065.
 PR 27-AUG-1999; 990S-0151066.
 PR 27-AUG-1999; 990S-0151080.
 PR 30-AUG-1999; 990S-0151303.
 PR 31-AUG-1999; 990S-0151438.
 PR 01-SEP-1999; 990S-0152364.
 PR 07-SEP-1999; 990S-0153070.
 PR 10-SEP-1999; 990S-0153758.
 PR 13-SEP-1999; 990S-0154018.
 PR 16-SEP-1999; 990S-0154039.
 PR 20-SEP-1999; 990S-0154779.
 PR 22-SEP-1999; 990S-0155139.
 PR 23-SEP-1999; 990S-0155486.
 PR 24-SEP-1999; 990S-0155659.
 PR 28-SEP-1999; 990S-0156458.
 PR 29-SEP-1999; 990S-0156596.
 PR 04-OCT-1999; 990S-0157117.
 PR 05-OCT-1999; 990S-0157753.
 PR 06-OCT-1999; 990S-0157865.
 PR 07-OCT-1999; 990S-0158029.
 PR 08-OCT-1999; 990S-0158232.
 PR 12-OCT-1999; 990S-0158369.
 PR 13-OCT-1999; 990S-0158294.
 PR 13-OCT-1999; 990S-0159295.
 PR 14-OCT-1999; 990S-0159329.
 PR 14-OCT-1999; 990S-0159330.
 PR 14-OCT-1999; 990S-0159331.
 PR 14-OCT-1999; 990S-0159637.
 PR 14-OCT-1999; 990S-0159636.
 PR 18-OCT-1999; 990S-0159584.
 PR 21-OCT-1999; 990S-0160741.
 PR 21-OCT-1999; 990S-0160767.
 PR 21-OCT-1999; 990S-0160768.
 PR 21-OCT-1999; 990S-0160770.
 PR 21-OCT-1999; 990S-0160814.
 PR 21-OCT-1999; 990S-0160815.
 PR 22-OCT-1999; 990S-0160980.
 PR 22-OCT-1999; 990S-0160981.
 PR 22-OCT-1999; 990S-0160989.
 PR 25-OCT-1999; 990S-0161404.
 PR 25-OCT-1999; 990S-0161405.
 PR 25-OCT-1999; 990S-0161406.
 PR 26-OCT-1999; 990S-0161359.
 PR 26-OCT-1999; 990S-0161360.
 PR 26-OCT-1999; 990S-0161361.
 PR 28-OCT-1999; 990S-0161920.
 PR 28-OCT-1999; 990S-0161992.
 PR 28-OCT-1999; 990S-0161993.
 PR 29-OCT-1999; 990S-0162142.

Query Match 0 68; Score 60; DB 21; Length 266;
 Best Local Similarity 28.68; Pred. No. 95;
 Matches 14; Conservative 11; Mismatches 24; Indels 0; Gaps 0;


```

PR 12 JUL 1999 0000 01420050
PR 13 JUL 1999 0000 01420050
PR 14 JUL 1999 0000 01420050
PR 15 JUL 1999 0000 01420050
PR 16 JUL 1999 0000 01420050
PR 17 JUL 1999 0000 01420050
PR 18 JUL 1999 0000 01420050
PR 19 JUL 1999 0000 01420050
PR 20 JUL 1999 0000 01420050
PR 21 JUL 1999 0000 01420050
PR 22 JUL 1999 0000 01420050
PR 23 JUL 1999 0000 01420050
PR 24 JUL 1999 0000 01420050
PR 25 JUL 1999 0000 01420050
PR 26 JUL 1999 0000 01420050
PR 27 JUL 1999 0000 01420050
PR 28 JUL 1999 0000 01420050
PR 29 JUL 1999 0000 01420050
PR 30 JUL 1999 0000 01420050
PR 31 JUL 1999 0000 01420050
PR 01 AUG 1999 0000 01420050
PR 02 AUG 1999 0000 01420050
PR 03 AUG 1999 0000 01420050
PR 04 AUG 1999 0000 01420050
PR 05 AUG 1999 0000 01420050
PR 06 AUG 1999 0000 01420050
PR 07 AUG 1999 0000 01420050
PR 08 AUG 1999 0000 01420050
PR 09 AUG 1999 0000 01420050
PR 10 AUG 1999 0000 01420050
PR 11 AUG 1999 0000 01420050
PR 12 AUG 1999 0000 01420050
PR 13 AUG 1999 0000 01420050
PR 14 AUG 1999 0000 01420050
PR 15 AUG 1999 0000 01420050
PR 16 AUG 1999 0000 01420050
PR 17 AUG 1999 0000 01420050
PR 18 AUG 1999 0000 01420050
PR 19 AUG 1999 0000 01420050
PR 20 AUG 1999 0000 01420050
PR 21 AUG 1999 0000 01420050
PR 22 AUG 1999 0000 01420050
PR 23 AUG 1999 0000 01420050
PR 24 AUG 1999 0000 01420050
PR 25 AUG 1999 0000 01420050
PR 26 AUG 1999 0000 01420050
PR 27 AUG 1999 0000 01420050
PR 28 AUG 1999 0000 01420050
PR 29 AUG 1999 0000 01420050
PR 30 AUG 1999 0000 01420050
PR 31 AUG 1999 0000 01420050
PR 01 SEP 1999 0000 01420050
PR 02 SEP 1999 0000 01420050
PR 03 SEP 1999 0000 01420050
PR 04 SEP 1999 0000 01420050
PR 05 SEP 1999 0000 01420050
PR 06 SEP 1999 0000 01420050
PR 07 SEP 1999 0000 01420050
PR 08 SEP 1999 0000 01420050
PR 09 SEP 1999 0000 01420050
PR 10 SEP 1999 0000 01420050
PR 11 SEP 1999 0000 01420050
PR 12 SEP 1999 0000 01420050
PR 13 SEP 1999 0000 01420050
PR 14 SEP 1999 0000 01420050
PR 15 SEP 1999 0000 01420050
PR 16 SEP 1999 0000 01420050
PR 17 SEP 1999 0000 01420050
PR 18 SEP 1999 0000 01420050
PR 19 SEP 1999 0000 01420050
PR 20 SEP 1999 0000 01420050
PR 21 SEP 1999 0000 01420050
PR 22 SEP 1999 0000 01420050
PR 23 SEP 1999 0000 01420050
PR 24 SEP 1999 0000 01420050
PR 25 SEP 1999 0000 01420050
PR 26 SEP 1999 0000 01420050
PR 27 SEP 1999 0000 01420050
PR 28 SEP 1999 0000 01420050
PR 29 SEP 1999 0000 01420050
PR 30 SEP 1999 0000 01420050
PR 01 OCT 1999 0000 01420050
PR 02 OCT 1999 0000 01420050
PR 03 OCT 1999 0000 01420050
PR 04 OCT 1999 0000 01420050
PR 05 OCT 1999 0000 01420050
PR 06 OCT 1999 0000 01420050
PR 07 OCT 1999 0000 01420050
PR 08 OCT 1999 0000 01420050
PR 09 OCT 1999 0000 01420050
PR 10 OCT 1999 0000 01420050
PR 11 OCT 1999 0000 01420050
PR 12 OCT 1999 0000 01420050
PR 13 OCT 1999 0000 01420050
PR 14 OCT 1999 0000 01420050
PR 15 OCT 1999 0000 01420050
PR 16 OCT 1999 0000 01420050
PR 17 OCT 1999 0000 01420050
PR 18 OCT 1999 0000 01420050
PR 19 OCT 1999 0000 01420050
PR 20 OCT 1999 0000 01420050
PR 21 OCT 1999 0000 01420050
PR 22 OCT 1999 0000 01420050
PR 23 OCT 1999 0000 01420050
PR 24 OCT 1999 0000 01420050
PR 25 OCT 1999 0000 01420050
PR 26 OCT 1999 0000 01420050
PR 27 OCT 1999 0000 01420050
PR 28 OCT 1999 0000 01420050
PR 29 OCT 1999 0000 01420050
PR 30 OCT 1999 0000 01420050
PR 31 OCT 1999 0000 01420050
PR 01 NOV 1999 0000 01420050
PR 02 NOV 1999 0000 01420050
PR 03 NOV 1999 0000 01420050
PR 04 NOV 1999 0000 01420050
PR 05 NOV 1999 0000 01420050
PR 06 NOV 1999 0000 01420050
PR 07 NOV 1999 0000 01420050
PR 08 NOV 1999 0000 01420050
PR 09 NOV 1999 0000 01420050
PR 10 NOV 1999 0000 01420050
PR 11 NOV 1999 0000 01420050
PR 12 NOV 1999 0000 01420050
PR 13 NOV 1999 0000 01420050
PR 14 NOV 1999 0000 01420050
PR 15 NOV 1999 0000 01420050
PR 16 NOV 1999 0000 01420050
PR 17 NOV 1999 0000 01420050
PR 18 NOV 1999 0000 01420050
PR 19 NOV 1999 0000 01420050
PR 20 NOV 1999 0000 01420050
PR 21 NOV 1999 0000 01420050
PR 22 NOV 1999 0000 01420050
PR 23 NOV 1999 0000 01420050
PR 24 NOV 1999 0000 01420050
PR 25 NOV 1999 0000 01420050
PR 26 NOV 1999 0000 01420050
PR 27 NOV 1999 0000 01420050
PR 28 NOV 1999 0000 01420050
PR 29 NOV 1999 0000 01420050
PR 30 NOV 1999 0000 01420050
PR 01 DEC 1999 0000 01420050
PR 02 DEC 1999 0000 01420050
PR 03 DEC 1999 0000 01420050
PR 04 DEC 1999 0000 01420050
PR 05 DEC 1999 0000 01420050
PR 06 DEC 1999 0000 01420050
PR 07 DEC 1999 0000 01420050
PR 08 DEC 1999 0000 01420050
PR 09 DEC 1999 0000 01420050
PR 10 DEC 1999 0000 01420050
PR 11 DEC 1999 0000 01420050
PR 12 DEC 1999 0000 01420050
PR 13 DEC 1999 0000 01420050
PR 14 DEC 1999 0000 01420050
PR 15 DEC 1999 0000 01420050
PR 16 DEC 1999 0000 01420050
PR 17 DEC 1999 0000 01420050
PR 18 DEC 1999 0000 01420050
PR 19 DEC 1999 0000 01420050
PR 20 DEC 1999 0000 01420050
PR 21 DEC 1999 0000 01420050
PR 22 DEC 1999 0000 01420050
PR 23 DEC 1999 0000 01420050
PR 24 DEC 1999 0000 01420050
PR 25 DEC 1999 0000 01420050
PR 26 DEC 1999 0000 01420050
PR 27 DEC 1999 0000 01420050
PR 28 DEC 1999 0000 01420050
PR 29 DEC 1999 0000 01420050
PR 30 DEC 1999 0000 01420050
PR 31 DEC 1999 0000 01420050

```

Query Match: 0.66; Score 61; DB 21; Length 699

Host: Local Similarity: 42.4%; Prod. No. 1100021

Matches: 14; Conservative: 5; Mismatches: 15; Gaps: 0

QY 1431 LSWPRTVTFEFLVPELALVAFVFFFESEEEQ 1663

16 328 Ldysqltqaaqkllppllqarssllkssssssss

RESULT: 73

AAR50653

1D AAR50653 standard; Protein: 731 AA.

XX AC AAR50653;

XX DL 19-MAR-2001 (first entry)

XX DE C. elegans spectrin beta chain/todrin protein SEQ ID No:14.

XX KW Human; Caenorhabditis elegans; RNC 5; splice variant; tomato-like worm;

XX KW protein-protein interaction; 140nt insertion.

XX CS Caenorhabditis elegans.

XX FN W0200073328-A2.

XX PU 07-DEC-2000.

XX PF 02-JUN-2000; 2000W00EP05108.

XX XX 01-JUN-1999; 99GB-0012755.

XX FA (FVG-) DEV:EN NV.

PR	01-SEP-1999	990S-0151940.	XX	06-SEP-2000.	
PR	07-SEP-1999	990S-0152363.	PD		
PR	10-SEP-1999	990S-0153070.	XX	25-FEB-2000; 2000EP-03C1439.	
PR	14-SEP-1999	990S-0153758.	XX		
PR	15-SEP-1999	990S-0154018.	XX	25-FEB-1999;	990S-0121825.
PR	16-SEP-1999	990S-0154039.	PR	05-MAR-1999;	990S-0123180.
PR	20-SEP-1999	990S-0154779.	PR	04-MAR-1999;	990S-0123548.
PR	22-SEP-1999	990S-0155139.	PR	23-MAR-1999;	990S-0125788.
PR	24-SEP-1999	990S-0155486.	PR	25-MAR-1999;	990S-0126264.
PR	24-SEP-1999	990S-0155659.	PR	29-MAR-1999;	990S-0126785.
PR	28-SEP-1999	990S-0156458.	PR	01-APR-1999;	990S-0127462.
PR	29-SEP-1999	990S-0156596.	PR	06-APR-1999;	990S-0128234.
PR	04-OCT-1999	990S-0157117.	PR	08-APR-1999;	990S-0128714.
PR	05-OCT-1999	990S-0157753.	PR	16-APR-1999;	990S-0129845.
PR	06-OCT-1999	990S-0157865.	PR	19-APR-1999;	990S-0130077.
PR	07-OCT-1999	990S-0158029.	PR	21-APR-1999;	990S-0130449.
PR	08-OCT-1999	990S-0158242.	PR	23-APR-1999;	990S-0130510.
PR	12-OCT-1999	990S-0158369.	PR	23-APR-1999;	990S-0130891.
PR	13-OCT-1999	990S-0159293.	PR	28-APR-1999;	990S-0131449.
PR	13-OCT-1999	990S-0159294.	PR	30-APR-1999;	990S-0132048.
PR	13-OCT-1999	990S-0159295.	PR	30-APR-1999;	990S-0132407.
PR	14-OCT-1999	990S-0159329.	PR	04-MAY-1999;	990S-0132484.
PR	14-OCT-1999	990S-0159340.	PR	05-MAY-1999;	990S-0132485.
PR	14-OCT-1999	990S-0159341.	PR	06-MAY-1999;	990S-0132486.
PR	14-OCT-1999	990S-0159637.	PR	06-MAY-1999;	990S-0132487.
PR	14-OCT-1999	990S-0159638.	PR	07-MAY-1999;	990S-0132863.
PR	18-OCT-1999	990S-0159584.	PR	11-MAY-1999;	990S-0134256.
PR	21-OCT-1999	990S-0160741.	PR	14-MAY-1999;	990S-0134218.
PR	21-OCT-1999	990S-0160767.	PR	14-MAY-1999;	990S-0134219.
PR	21-OCT-1999	990S-0160768.	PR	14-MAY-1999;	990S-0134221.
PR	21-OCT-1999	990S-0160770.	PR	14-MAY-1999;	990S-0134370.
PR	21-OCT-1999	990S-0160814.	PR	18-MAY-1999;	990S-0134768.
PR	21-OCT-1999	990S-0160815.	PR	19-MAY-1999;	990S-0134941.
PR	22-OCT-1999	990S-0160980.	PR	20-MAY-1999;	990S-0135124.
PR	22-OCT-1999	990S-0160981.	PR	21-MAY-1999;	990S-0135353.
PR	22-OCT-1999	990S-0160989.	PR	24-MAY-1999;	990S-0135629.
PR	25-OCT-1999	990S-0161404.	PR	25-MAY-1999;	990S-0136021.
PR	25-OCT-1999	990S-0161405.	PR	27-MAY-1999;	990S-0136392.
PR	25-OCT-1999	990S-0161406.	PR	28-MAY-1999;	990S-0136782.
PR	26-OCT-1999	990S-0161359.	PR	01-JUN-1999;	990S-0137222.
PR	26-OCT-1999	990S-0161360.	PR	03-JUN-1999;	990S-0137528.
PR	26-OCT-1999	990S-0161361.	PR	04-JUN-1999;	990S-0137502.
PR	28-OCT-1999	990S-0161920.	PR	07-JUN-1999;	990S-0137724.
PR	28-OCT-1999	990S-0161992.	PR	08-JUN-1999;	990S-0138094.
PR	28-OCT-1999	990S-0161993.	PR	10-JUN-1999;	990S-0138540.
PR	29-OCT-1999	990S-0162142.	PR	10-JUN-1999;	990S-0138847.
PR			PR	14-JUN-1999;	990S-0139119.
PR			PR	16-JUN-1999;	990S-0139452.
PR			PR	16-JUN-1999;	990S-0139453.
PR			PR	17-JUN-1999;	990S-0139492.
PR			PR	18-JUN-1999;	990S-0139454.
PR			PR	18-JUN-1999;	990S-0139455.
PR			PR	18-JUN-1999;	990S-0139456.
PR			PR	18-JUN-1999;	990S-0139457.
PR			PR	18-JUN-1999;	990S-0139458.
PR			PR	18-JUN-1999;	990S-0139459.
PR			PR	18-JUN-1999;	990S-0139460.
PR			PR	18-JUN-1999;	990S-0139461.
PR			PR	18-JUN-1999;	990S-0139462.
PR			PR	18-JUN-1999;	990S-0139463.
PR			PR	18-JUN-1999;	990S-0139750.
PR			PR	18-JUN-1999;	990S-0139763.
PR			PR	21-JUN-1999;	990S-0139817.
PR			PR	21-JUN-1999;	990S-0139899.
PR			PR	23-JUN-1999;	990S-0140353.
PR			PR	23-JUN-1999;	990S-0140354.
PR			PR	24-JUN-1999;	990S-0140695.
PR			PR	28-JUN-1999;	990S-0140823.
PR			PR	29-JUN-1999;	990S-0140991.
PR			PR	30-JUN-1999;	990S-0141287.
PR			PR	01-JUL-1999;	990S-0141842.
PR			PR	01-JUL-1999;	990S-0142154.

Query Match 0.68; Score 61; DB 21; Length 369;

Best Local Similarity 42.54; Pred. Re. 1.1e-02;

Match 15; Conservative 5; Missat-Per 15; Indels 0; Gaps 0;

07 1631 15WRRKIVTFFKIVMELIAIVGPPKIVGTFQ 1669

DB 329 1478:qlvvtatqllpvltrvreslrvccccc 360

RESULT 72

AA039086

ID AA039086 standard; Protein; 369 AA.

AC AA039086;

XX

XX

DT 18-OCT-2000 (first entry)

DE

Arabidopsis thaliana protein fragment SEQ ID NO: 62864.

XX

XX

KW protein identification, signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX termination sequence.

OS Arabidopsis thaliana.

XX

PR EF103405 At.

KW PR 24 JUN 1999; 990S-0140854.
 KW PR 24 JUN 1999; 990S-0140695.
 KW PR 28 JUN 1999; 990S-0140823.
 XX PR 29 JUN 1999; 990S-0140991.
 XX PR 30 JUN 1999; 990S-0141287.
 XX PR 01 JUL 1999; 990S-0141842.
 XX PR 01 JUL 1999; 990S-0142154.
 XX PR 02 JUL 1999; 990S-0142055.
 XX PR 06 JUL 1999; 990S-0142490.
 XX PR 06 JUL 1999; 990S-0142803.
 XX PR 09 JUL 1999; 990S-0142920.
 XX PR 12 JUL 1999; 990S-0142977.
 XX PR 14 JUL 1999; 990S-0143542.
 XX PR 15 JUL 1999; 990S-0143624.
 XX PR 15 JUL 1999; 990S-0143605.
 XX PR 16 JUL 1999; 990S-0144085.
 XX PR 16 JUL 1999; 990S-0144086.
 XX PR 19 JUL 1999; 990S-0144325.
 XX PR 19 JUL 1999; 990S-0144331.
 XX PR 19 JUL 1999; 990S-0144332.
 XX PR 19 JUL 1999; 990S-0144333.
 XX PR 19 JUL 1999; 990S-0144334.
 XX PR 21 JUL 1999; 990S-0144344.
 XX PR 21 JUL 1999; 990S-0145086.
 XX PR 21 JUL 1999; 990S-0145088.
 XX PR 22 JUL 1999; 990S-0145085.
 XX PR 22 JUL 1999; 990S-0145087.
 XX PR 22 JUL 1999; 990S-0145089.
 XX PR 22 JUL 1999; 990S-0145132.
 XX PR 23 JUL 1999; 990S-0145145.
 XX PR 24 JUL 1999; 990S-0145218.
 XX PR 24 JUL 1999; 990S-0145224.
 XX PR 25 JUL 1999; 990S-0145276.
 XX PR 27 JUL 1999; 990S-0145913.
 XX PR 27 JUL 1999; 990S-0145918.
 XX PR 27 JUL 1999; 990S-0145919.
 XX PR 28 JUL 1999; 990S-0145951.
 XX PR 02 AUG 1999; 990S-0146386.
 XX PR 02 AUG 1999; 990S-0146488.
 XX PR 02 AUG 1999; 990S-0146489.
 XX PR 04 AUG 1999; 990S-0147038.
 XX PR 04 AUG 1999; 990S-0147204.
 XX PR 04 AUG 1999; 990S-0147302.
 XX PR 05 AUG 1999; 990S-0147192.
 XX PR 05 AUG 1999; 990S-0147260.
 XX PR 06 AUG 1999; 990S-0147303.
 XX PR 06 AUG 1999; 990S-0147416.
 XX PR 09 AUG 1999; 990S-0147493.
 XX PR 09 AUG 1999; 990S-0147935.
 XX PR 10 AUG 1999; 990S-0148171.
 XX PR 11 AUG 1999; 990S-0148319.
 XX PR 12 AUG 1999; 990S-0148441.
 XX PR 13 AUG 1999; 990S-0148565.
 XX PR 13 AUG 1999; 990S-0148684.
 XX PR 16 AUG 1999; 990S-0149688.
 XX PR 17 AUG 1999; 990S-0149175.
 XX PR 18 AUG 1999; 990S-0149426.
 XX PR 20 AUG 1999; 990S-0149722.
 XX PR 20 AUG 1999; 990S-0149723.
 XX PR 20 AUG 1999; 990S-0149929.
 XX PR 23 AUG 1999; 990S-0149902.
 XX PR 23 AUG 1999; 990S-0149930.
 XX PR 25 AUG 1999; 990S-0150566.
 XX PR 26 AUG 1999; 990S-0150884.
 XX PR 27 AUG 1999; 990S-0151065.
 XX PR 27 AUG 1999; 990S-0151066.
 XX PR 27 AUG 1999; 990S-0151080.
 XX PR 30 AUG 1999; 990S-0151303.
 XX PR 31 AUG 1999; 990S-0151438.

XX PR 24 JUN 1999; 990S-0142823.
 XX PR 24 JUN 1999; 990S-0143198.
 XX PR 25 MAR 1999; 990S-0143544.
 XX PR 25 MAR 1999; 990S-0143696.
 XX PR 25 MAR 1999; 990S-0143944.
 XX PR 25 MAR 1999; 990S-0143945.
 XX PR 25 MAR 1999; 990S-0143946.
 XX PR 25 MAR 1999; 990S-0143947.
 XX PR 25 MAR 1999; 990S-0143948.
 XX PR 25 MAR 1999; 990S-0143949.
 XX PR 25 MAR 1999; 990S-0143950.
 XX PR 25 MAR 1999; 990S-0143951.
 XX PR 25 MAR 1999; 990S-0143952.
 XX PR 25 MAR 1999; 990S-0143953.
 XX PR 25 MAR 1999; 990S-0143954.
 XX PR 25 MAR 1999; 990S-0143955.
 XX PR 25 MAR 1999; 990S-0143956.
 XX PR 25 MAR 1999; 990S-0143957.
 XX PR 25 MAR 1999; 990S-0143958.
 XX PR 25 MAR 1999; 990S-0143959.
 XX PR 25 MAR 1999; 990S-0143960.
 XX PR 25 MAR 1999; 990S-0143961.
 XX PR 25 MAR 1999; 990S-0143962.
 XX PR 25 MAR 1999; 990S-0143963.
 XX PR 25 MAR 1999; 990S-0143964.
 XX PR 25 MAR 1999; 990S-0143965.
 XX PR 25 MAR 1999; 990S-0143966.
 XX PR 25 MAR 1999; 990S-0143967.
 XX PR 25 MAR 1999; 990S-0143968.
 XX PR 25 MAR 1999; 990S-0143969.
 XX PR 25 MAR 1999; 990S-0143970.
 XX PR 25 MAR 1999; 990S-0143971.
 XX PR 25 MAR 1999; 990S-0143972.
 XX PR 25 MAR 1999; 990S-0143973.
 XX PR 25 MAR 1999; 990S-0143974.
 XX PR 25 MAR 1999; 990S-0143975.
 XX PR 25 MAR 1999; 990S-0143976.
 XX PR 25 MAR 1999; 990S-0143977.
 XX PR 25 MAR 1999; 990S-0143978.
 XX PR 25 MAR 1999; 990S-0143979.
 XX PR 25 MAR 1999; 990S-0143980.
 XX PR 25 MAR 1999; 990S-0143981.
 XX PR 25 MAR 1999; 990S-0143982.
 XX PR 25 MAR 1999; 990S-0143983.
 XX PR 25 MAR 1999; 990S-0143984.
 XX PR 25 MAR 1999; 990S-0143985.
 XX PR 25 MAR 1999; 990S-0143986.
 XX PR 25 MAR 1999; 990S-0143987.
 XX PR 25 MAR 1999; 990S-0143988.
 XX PR 25 MAR 1999; 990S-0143989.
 XX PR 25 MAR 1999; 990S-0143990.
 XX PR 25 MAR 1999; 990S-0143991.
 XX PR 25 MAR 1999; 990S-0143992.
 XX PR 25 MAR 1999; 990S-0143993.
 XX PR 25 MAR 1999; 990S-0143994.
 XX PR 25 MAR 1999; 990S-0143995.
 XX PR 25 MAR 1999; 990S-0143996.
 XX PR 25 MAR 1999; 990S-0143997.
 XX PR 25 MAR 1999; 990S-0143998.
 XX PR 25 MAR 1999; 990S-0143999.
 XX PR 25 MAR 1999; 990S-0144000.

```

PR 16-JUN-1999; 99US-0139462.
PR 16-JUN-1999; 99US-0139464.
PR 18-JUN-1999; 99US-0139760.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140605.
PR 24-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 29-JUN-1999; 99US-0141287.
PR 30-JUN-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 01-JUL-1999; 99US-0142055.
PR 02-JUL-1999; 99US-0142390.
PR 06-JUL-1999; 99US-0142861.
PR 08-JUL-1999; 99US-0142920.
PR 09-JUL-1999; 99US-0142977.
PR 12-JUL-1999; 99US-0143542.
PR 13-JUL-1999; 99US-0143624.
PR 14-JUL-1999; 99US-0144005.
PR 15-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144542.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 23-JUL-1999; 99US-0145276.
PR 26-JUL-1999; 99US-0145213.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 27-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146380.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147046.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.

```

```

PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0158029.
PR 07-OCT-1999; 99US-0158232.
PR 08-OCT-1999; 99US-0158369.
PR 12-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 0.68; Score 61; DB 21; Length 357;
Best Local Similarity 42.48; Prod. No. 11c-02;
Matches 14; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

```

```

OY 1631 ISWKKTIVIRFKLVDPDLAIYQKKKKGHEFQ 1663
DB 316 IYKRAIYTTPLIPIIIIAIESIPETETETET 318

```

RESULT 71

AAG18559

ID AAG18559 standard; Protein; 369 AA.

XX AAG18559;

XX

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 20017.

XX

PR	17 AUG 1999	9908-014025	XX	AA04-0089;
PR	18 AUG 1999	9908-014046	AC	
PR	20 AUG 1999	9908-014022	XX	18 OCT-2000 (first entry)
PR	21 AUG 1999	9908-014023	XX	
PR	23 AUG 1999	9908-014026	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 6205.
PR	25 AUG 1999	9908-014027	XX	Protein identification: signal transduction pathway; metabolic pathway;
PR	26 AUG 1999	9908-014030	KW	hybridisation assay; detection; map; post; gene expression; control; promoter;
PR	27 AUG 1999	9908-014031	KW	termination sequence.
PR	28 AUG 1999	9908-014034	XX	Arabidopsis thaliana.
PR	29 AUG 1999	9908-014035	XX	
PR	30 AUG 1999	9908-014036	PN	EP1043405-AZ.
PR	31 AUG 1999	9908-014037	PD	06-SEP-2000.
PR	32 AUG 1999	9908-014038	XX	
PR	33 AUG 1999	9908-014039	XX	25 FEB-2000; 2000EP-040149.
PR	34 AUG 1999	9908-014040	XX	
PR	35 AUG 1999	9908-014041	PR	25-FEB-1999; 9908-0121825.
PR	36 AUG 1999	9908-014042	PR	05-MAR-1999; 9908-0124180.
PR	37 AUG 1999	9908-014043	PR	09-MAR-1999; 9908-0124548.
PR	38 AUG 1999	9908-014044	PR	24-MAR-1999; 9908-0125788.
PR	39 AUG 1999	9908-014045	PR	25-MAR-1999; 9908-0126264.
PR	40 AUG 1999	9908-014046	PR	29-MAR-1999; 9908-0126785.
PR	41 AUG 1999	9908-014047	PR	01-APR-1999; 9908-0127462.
PR	42 AUG 1999	9908-014048	PR	06-APR-1999; 9908-0128234.
PR	43 AUG 1999	9908-014049	PR	08-APR-1999; 9908-0128714.
PR	44 AUG 1999	9908-014050	PR	16-APR-1999; 9908-0129845.
PR	45 AUG 1999	9908-014051	PR	19-APR-1999; 9908-0130077.
PR	46 AUG 1999	9908-014052	PR	21-APR-1999; 9908-0130449.
PR	47 AUG 1999	9908-014053	PR	24-APR-1999; 9908-0130510.
PR	48 AUG 1999	9908-014054	PR	28-APR-1999; 9908-0130801.
PR	49 AUG 1999	9908-014055	PR	30-APR-1999; 9908-0131449.
PR	50 AUG 1999	9908-014056	PR	04-APR-1999; 9908-0132048.
PR	51 AUG 1999	9908-014057	PR	08-MAY-1999; 9908-0132407.
PR	52 AUG 1999	9908-014058	PR	05-MAY-1999; 9908-0132484.
PR	53 AUG 1999	9908-014059	PR	06-MAY-1999; 9908-0132485.
PR	54 AUG 1999	9908-014060	PR	07-MAY-1999; 9908-0132486.
PR	55 AUG 1999	9908-014061	PR	08-MAY-1999; 9908-0132487.
PR	56 AUG 1999	9908-014062	PR	09-MAY-1999; 9908-0132864.
PR	57 AUG 1999	9908-014063	PR	11-MAY-1999; 9908-0134256.
PR	58 AUG 1999	9908-014064	PR	14-MAY-1999; 9908-0134218.
PR	59 AUG 1999	9908-014065	PR	15-MAY-1999; 9908-0134219.
PR	60 AUG 1999	9908-014066	PR	16-MAY-1999; 9908-0134221.
PR	61 AUG 1999	9908-014067	PR	17-MAY-1999; 9908-0134370.
PR	62 AUG 1999	9908-014068	PR	18-MAY-1999; 9908-0134768.
PR	63 AUG 1999	9908-014069	PR	19-MAY-1999; 9908-0134941.
PR	64 AUG 1999	9908-014070	PR	20-MAY-1999; 9908-0135124.
PR	65 AUG 1999	9908-014071	PR	21-MAY-1999; 9908-0135353.
PR	66 AUG 1999	9908-014072	PR	24-MAY-1999; 9908-0135629.
PR	67 AUG 1999	9908-014073	PR	25-MAY-1999; 9908-0136021.
PR	68 AUG 1999	9908-014074	PR	26-MAY-1999; 9908-0136392.
PR	69 AUG 1999	9908-014075	PR	27-MAY-1999; 9908-0136782.
PR	70 AUG 1999	9908-014076	PR	28-MAY-1999; 9908-0137222.
PR	71 AUG 1999	9908-014077	PR	01-JUN-1999; 9908-0137528.
PR	72 AUG 1999	9908-014078	PR	04-JUN-1999; 9908-0137502.
PR	73 AUG 1999	9908-014079	PR	07-JUN-1999; 9908-0137724.
PR	74 AUG 1999	9908-014080	PR	08-JUN-1999; 9908-0138094.
PR	75 AUG 1999	9908-014081	PR	10-JUN-1999; 9908-0138540.
PR	76 AUG 1999	9908-014082	PR	14-JUN-1999; 9908-0148847.
PR	77 AUG 1999	9908-014083	PR	16-JUN-1999; 9908-0139119.
PR	78 AUG 1999	9908-014084	PR	17-JUN-1999; 9908-0139454.
PR	79 AUG 1999	9908-014085	PR	18-JUN-1999; 9908-0139455.
PR	80 AUG 1999	9908-014086	PR	18-JUN-1999; 9908-0139456.
PR	81 AUG 1999	9908-014087	PR	18-JUN-1999; 9908-0139457.
PR	82 AUG 1999	9908-014088	PR	18-JUN-1999; 9908-0139458.
PR	83 AUG 1999	9908-014089	PR	18-JUN-1999; 9908-0139459.
PR	84 AUG 1999	9908-014090	PR	18-JUN-1999; 9908-0139460.
PR	85 AUG 1999	9908-014091	PR	18-JUN-1999; 9908-0139461.

2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37

2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37

2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37

2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37

2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37

2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37

2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37

2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37

2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37

2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37

2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37

2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37 2001-08-08 15:42:37


```
PR 04-AUG-1999; 99US-0147234.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147304.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149940.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154053.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155579.
PR 26-SEP-1999; 99US-0156358.
PR 26-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159339.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160769.
PR 21-OCT-1999; 99US-0160769.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.

PR 29-OCT-1999; 99US-0162142.

Query Match: 0 68; Score 61; DB 21; Length 152;
Best Local Similarity 42.48; Pred. No. 1e-02;
Matches 14; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

Qy 1631 ISWKKTIVTPELKLVPDLIAIVQPKKKKEFEQ 1663
Db 311 IYKRGIVYTKPIPIIITAVRESIVQVQVQVQV 343

RESULT 68
AAG49090
ID AAG49090 standard; Protein; 352 AA.
XX
AC AAG49090;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 62066.
XX
KW Protein identification, signal transduction pathway, metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000
XX
PE 25-FEB-2000; 2000APP-0301439.
XX
PE 25-FEB-1999; 99US-0121925.
PR 05-MAR-1999; 99US-0121980.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
```


CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
 CC e.g. cerebral ischemia, aneurysms, nervous system disorders e.g.
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and
 CC ocular disorders e.g. corneal infection. The polypeptides can also be
 CC used to aid wound healing and epithelial cell proliferation, to prevent
 CC skin aging due to sunburn, to maintain organs before transplantation, for
 CC supporting cell culture of primary tissues, to regenerate tissues and in
 CC chemotaxis. The polypeptides can also be used as a food additive or
 CC preservative to increase or decrease storage capabilities, fat content,
 CC lipid, protein, carbohydrate, vitamins, minerals, cofactors and other
 CC nutritional components. AAF22297 to AAF22315 and AAB62048 represent
 CC sequences used in the exemplification of the present invention.

XX
 SQ Sequence 226 AA;

Query Match 0.6%; Score 62; DB 21; Length 226;
 Best Local Similarity 41.7%; Pred. No. 40;
 Matches 15; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 919 VVTSLLINLGSPVKEVRRRAAAGCLOALSCVASPPFYI 954
 ||| | | | | | | | | | | | | | | |
 DB 155 VTKLILSSSPSMARVIAALQMBALTRIPYVLL 190

RESULT 60

AAAY67644
 ID AAY67644 standard; Protein; 244 AA.

XX
 AC AAY67644;

XX
 DT 07-OCT-1999 (first entry)

XX
 DE Protein which is specific to Chlamydia trachomatis.

XX
 KW Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
 KW paratrachoma; inclusion conjunctivitis; genital disease; peritritis;
 KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
 KW Bartholinitis; pneumopathy; venereal lymphogranulomatosis.

XX
 CS Chlamydia trachomatis.

XX
 PN W09928475-A2.

XX
 PD 10-JUN-1999.

XX
 PF 27-NOV-1998; 98WG 1B01939.

XX
 PR 04-NOV-1998; 98US 0107077.

XX
 PR 28-NOV-1997; 97ER-0015041.

XX
 PR 17-DEC-1997; 97ER-0016034.

XX
 PA (GIST) GENSET.

XX
 PT Griffiths R;

XX
 DB WFL 1999 171125/31.

XX
 PT Genome sequence of Chlamydia trachomatis

XX
 PS Disclosure, Page 1273, 1755pp; English.

XX
 CC AAY67644 (3795) are encoded by open reading frames (ORFs) of the genome
 CC of Chlamydia trachomatis (see AAY67644). The polypeptides can be used as
 CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
 CC can also be used to control growth of the microorganism. Chlamydia
 CC trachomatis is responsible for a large number of diseases, e.g. eye
 CC diseases such as conventional trachoma, nonendemic trachoma,
 CC paratrachoma, and inclusion conjunctivitis; genital diseases such as
 CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis,
 CC peritritis, Bartholinitis, pneumopathy in breast feeding infants;
 CC and venereal lymphogranulomatosis. The polypeptides of the invention
 CC may be of use in treating these diseases.

XX
 SQ Sequence 234 AA;

Query Match 0.6%; Score 62; DB 20; Length 244;
 Best Local Similarity 42.9%; Pred. No. 42;
 Matches 12; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 1018 IMKVIQGVNCRMVLSQLPMAHQLEKI 1045
 ||| | | | | | | | | | | | | | | |
 DB 72 IMDIQGLSGLIQQIGLSVLEEL 99

RESULT 61

AAB61304
 ID AAB61304 standard; Protein; 615 AA.

XX
 AC AAB61304;

XX
 DT 30-MAR-2001 (first entry)

XX
 DE Human transcriptional regulator protein #4.

XX
 KW Human; transcriptional regulator protein; IXPEG.

XX
 CS Homo sapiens.

XX
 PN W0200079954-A2.

XX
 PD 29-DEC-2000.

XX
 PF 15-JUN-2000; 2000WO-0516766

XX
 PR 18-JUN-1999; 99US-0140109.

XX
 PA (INCY-) INCYTE GEN-MICS INC.

XX
 PI Lal P, Yue H, Tang YT, Baughn ME, Azimzai Y, Tran B;

XX
 DR WPI; 2001-041425/05.

XX
 PT Isolated polypeptide with a human transcriptional regulator protein
 PT sequence is useful for the diagnosis, prevention and treatment of
 PT disorders associated with the immune, reproductive and cardiovascular
 PT systems -

XX
 PS Claim 1; Page 95-96; 142pp; English.

XX
 CC The present invention relates to human transcriptional regulator
 CC protein (IXPEG) sequences. The antagonist and an agonist of the proteins
 CC of the invention are used to treat disorders associated with decreased
 CC or increased expression of activity of IXPEG.

XX
 SQ Sequence 615 AA;

Query Match 0.6%; Score 62; DB 22; Length 615;
 Best Local Similarity 41.7%; Pred. No. 170402;
 Matches 15; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 919 VVTSLLINLGSPVKEVRRRAAAGCLOALSCVASPPFYI 954
 ||| | | | | | | | | | | | | | | |
 DB 540 VTKLILSSSPSMARVIAALQMBALTRIPYVLL 575

RESULT 62

AAAY69294
 ID AAY69294 standard; Protein; 701 AA.

XX
 AC AAY69294;

XX
 DT 19-JUN-2000 (first entry)

XX

XX Yeast GAL4/YMR117c fusion protein.
 XX Two hybrid assay, fusion, transcriptional activation, information, matrix.
 XX array; screening; reporter gene; identification.
 XX
 XX Synthetic.
 XX Saccharomyces cerevisiae.
 XX
 XX Key location/qualifiers
 XX Domain 1..147
 XX /note- "DNA binding domain of GAL4"
 XX Protein 148..378
 XX /note- "Yeast YMR117c protein"
 XX Mis-difference 378
 XX /note- "Encoded by TGA"
 XX
 XX W09935256-A1.
 XX
 XX 15-JUL-1999.
 XX
 XX 06-JAN-1999; 9909-IB00048.
 XX
 XX 06-JAN-1998; 9805-0003335.
 XX
 XX 17-SEP-1998; 9805-0154972.
 XX
 XX (IRF) and 7K1140K.
 XX (STRD) and 5148VAD.
 XX (AFFY-) AFFYMETRIX.
 XX
 XX Leclair P, Fromont-Karlme M, Cho R, Davis R, Lockhart D;
 XX Podicka L;
 XX
 XX WPI: 1999-590680/50.
 XX N FSLB; AAZ28176.
 XX
 XX screening polynucleotides with whole genome oligonucleotide or
 XX polynucleotide arrays -
 XX
 XX Example 1; Page 36-37; 41pp; English.
 XX
 XX This sequence represents a yeast GAL4/YMR117c fusion protein. YMR117c is
 XX a yeast GPF (open reading frame) which was found by two-hybrid analysis
 XX to interact with the U2 snRNP-associated splicing factor, Prp1p. This
 XX was used in a novel two hybrid screening method for the identification of
 XX nucleic acids by direct hybridisation to high-density oligonucleotide
 XX arrays. Two hybrid assays exploit the ability of a pair of interacting
 XX proteins to bind a transactivator domain into close proximity with a
 XX DNA-binding site which regulates the expression of a reporter gene. The
 XX assay employs chimeric genes which express two types of fusion protein.
 XX one hybrid protein, contains a transactivation domain (e.g., that of
 XX GAL4) fused to a first test protein (e.g., YMR117c protein). A second
 XX fusion protein contains a second test protein that potentially interacts
 XX with the first test protein. If the two test proteins interact, they
 XX bring the two domains of the transcriptional activator into close
 XX proximity. The transcriptional activator can then bind to the
 XX corresponding transcriptional activator binding site on the reporter
 XX gene, thus effecting reporter gene transcription, and causing a
 XX detectable signal to be produced. The two hybrid assay can be used to
 XX test a large number of proteins simultaneously to determine whether they
 XX interact with a particular protein. Such a two-hybrid assay was
 XX performed with a vector containing the yeast GAL4/YMR117c DNA sequence
 XX and the clones encoding interacting proteins were selected. The DNA
 XX inserts of the partner proteins were amplified and hybridised to an
 XX oligonucleotide probe matrix. Together, the probes encompassed all the
 XX coding sequences on the yeast genome, enabling the hybridisation
 XX locations of the genes encoding partner proteins to be identified. The
 XX methods of the invention thus allow large scale screening for DNA
 XX sequences that have functional utility and avoids the systematic
 XX sequencing of DNA inserts of interest that is required by prior art
 XX methods. The methods can be used to screen polynucleotides, especially
 XX those from a host organism whose genome has already been sequenced.

SQ Sequence 378 AA;

Query Match 0.6%; Score 67; DR 20; Length 378;
 Best Local Similarity 31.5%; Pred. NO. 62;
 Matches 17; Conservative 8; Mismatches 29; Indels 0; Gaps 0;

Q7 1430 AAAAGKDAIIFAVVPEWPSVQVHQIGUSLNNIIQYLKLPEEKETIPK 1483
 DB 162 aaamaqfndhdpvclakviesdihppkark;hhldhdkhkwssarisk 215

RESULT 53

AAB53398

ID AAB53398 standard; Protein; 752 AA.

XX

AC AAB53398;

XX

XX 09-MAR-2001 (first entry)

XX

DE Human colon cancer antigen protein sequence SEQ ID NO:938.

XX

KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW identification; cytostatic; cardioprotective; neuroprotective; vulnary;
 KW immunomodulatory; muscular; gynaecological; gastrointestinal;
 KW nephrotic; antineoplastic; antibacterial; gene therapy; wound;
 KW neural disease; immune system disorder; vascular disorder;
 KW reproductive disorder; gastrointestinal disorder; renal disorder;
 KW infectious disease; cardiovascular disorder.

XX

OS Homo sapiens.

XX

PN WC2000055351-A1.

XX

PD 21-SEP-2000.

XX

PF 08-MAR-2000; 2000WO-US05883.

XX

PR 12 MAR 1999; 9805-0124270

XX

XX (HUMA-) HUMAN GENOME SCI INC.

XX

XX Poser CA, Puben SM;

XX

XX WPI: 2600-087514/55.

XX

XX N-PSUR; AAC98155

XX

Colon cancer associated gene sequences, referred to as colon cancer
 antigens, useful for the treatment, prevention, and diagnosis of colon
 disorders such as colon cancer -

XX

XX Claim 11; Page 1495-1499; 2104pp; English.

XX

AAC9799; to AAC98763 encode the human colon cancer associated proteins,
 called human colon cancer antigens, given in AAB53234 to AAB54006. The
 human colon cancer antigens can have cytostatic, cardioprotective, muscular,
 neuroprotective, immunomodulatory, gynaecological, gastrointestinal, and
 vulnary, nephrotic, antineoplastic, antibacterial activities, and
 can be used in gene therapy. The colon cancer antigen polynucleotides,
 proteins and antibodies to the proteins are useful for the prevention,
 treatment and diagnosis of colon disorders, such as colon cancer. The
 polynucleotides may be used in diagnostics and research, such as for
 chromosome identification, and as hybridisation probes. The proteins
 may also be used to prevent diseases such as neural disorders, immune
 system disorders, muscular disorders, reproductive disorders,
 gastrointestinal disorders, wounds, renal disorders, infectious
 diseases, and cardiovascular disorders. AAC98764 to AAC98772 and
 AAB54007 represent sequences used in the exemplification of the present
 invention.

XX

XX Sequence 752 AA;

XX

SQ

CC 6. Rat rattus (rathsp70 - AAR03927);
 CC 7. Xenopus laevis (xl170 - AAR03928);
 CC 8. Homo sapiens (humisp70 - AAR03929);
 CC 9. Gallus gallus (chkhsp70 - AAR03930);
 CC 10. Zca mayis (mzhsp70 - AAR03931);
 CC 11. Serratia marcescens (smahsp70 - AAR03932).
 CC The proteins having homology to hsp's of T. cruzi can be used in
 CC vaccines and diagnosis involving e.g. Trypanosoma, Mycoplasma and
 CC Mycobacteria species.
 XX Sequence 605 AA;

Query Match 0.6%; Score 64; DB 11; Length 605;
 Best Local Similarity 32.7%; Pred. No. 89;
 Matches 17; Conservative 12; Mismatches 23; Indels 0; Gaps 0;

QY 548 FSSVFTISNLNLFORAELSKNGEWEVLKTAADILLKEILSENDSQV 599
 Db 514 tttkttkdlegkveevrtkancaakdalkaalekndicoikakkdelgolv 565

RESULT 43
 AAY22469
 ID AAY22469 standard; Protein, 626 AA.

XX AC AAY22469,
 XX DT 29-SEP-1999 (first entry)
 XX DE Human 6myc-N-sel-10 protein sequence.
 XX KW Sel 10, human, presenilin, Alzheimer's disease; FS-1; FS-2; hippocampus;
 XX RW therapy; mammary gland.

XX OS Homo sapiens.
 XX OS Synthetic.
 XX PN W09932623-A1.

XX PD 01-JUL-1999.
 XX PF 17-DEC-1998; GAWO-US26820
 XX PR 19-DEC-1997; 97US-0068243.
 XX PA (PHAA) PHARMACIA & UPJOHN CO.

XX PI Gurney ME, Li J, Pauley AM;
 XX DR WPI: 1999-458026/38.
 XX DR N-PSDB; AAX99712.

XX PT New isolated human sel-10 polypeptides
 XX PS Example 3; Page 74-77; 91pp; English.

XX CC This sequence is an epitope-tagged version of a human sel-10 protein
 CC of the invention. The human sel-10 proteins of the invention are isolated
 CC from hippocampus and mammary gland. The polypeptides can be used to
 CC alter presenilin function. Compounds which inhibit either the expression
 CC or the activity of the human sel-10 polypeptides may reverse the effects
 CC of mutations to presenilin-1 (PS-1) or PS-2, and therefore may be useful
 CC for the prevention or treatment of Alzheimer's disease.

XX Sequence 626 AA;

Query Match 0.6%; Score 64; DB 20; Length 626;
 Best Local Similarity 27.5%; Pred. No. 93;
 Matches 19; Conservative 16; Mismatches 40; Indels 0; Gaps 3.

QY 50 TGLELLGIDPSFEQFEAPLFSOLAKTLFESVQTKAVNKQLDENISLFIHLSPYFLKPK 109

Db 164 sgpckllaladelidsceptqvkummqvlepaqfgdrifslpkelalyvislfpckdlilqa 223

QY 110 AOKCLEWLI 118
 Db 224 aqterywri 232

RESULT 44
 AAB59201
 ID AAB59201 standard; Protein, 626 AA

XX AC AAB59201;
 XX DT 23-MAR-2001 (first entry)
 XX DE Protein encoded by construct 6myc-N-sel-10.

XX KW Sel-10; human; Alzheimer's disease; Abeta.
 XX OS Synthetic.
 XX PN W0200075328-A1.

XX PD 14-DEC-2000
 XX PF 23-MAY-2000; 2000W0-US09814.
 XX PR 09-JUN-1999; 99US-0328877.

XX PA (PHAA) PHARMACIA & UPJOHN CO
 XX PI Pauley AM, Gurney ME, Li J;
 XX DR WPI: 2001-102404/11.

XX CC New human sel-10 polypeptides and their encoding polynucleotides,
 XX PT useful for raising antibodies for detecting sel-10 polypeptide
 XX PT expression and as drug targets in the treatment of Alzheimer's disease
 XX PT .
 XX PS Example 3; Page 95-98, 116pp; English.

XX CC The present invention relates to human sel-10, the sel-10 proteins of
 CC the invention are useful for raising monoclonal or polyclonal
 CC antibodies useful in diagnostic assays for detecting sel-10
 CC polypeptide expression. The sel-10 polypeptides are also useful as drug
 CC targets for decreasing antibody levels in the treatment of Alzheimer's
 CC disease. It is also useful for identifying agents capable of
 CC altering the production level of Abeta. The polynucleotides are useful
 CC for developing assays for identifying agents capable of interfering
 CC with the biological pathways that lead to Alzheimer's disease.

XX Sequence 626 AA;

Query Match 0.6%; Score 64; DB 22; Length 626;
 Best Local Similarity 27.5%; Pred. No. 93;
 Matches 19; Conservative 10; Mismatches 40; Indels 0; Gaps 0;

QY 50 TGLELLGIDPSFEQFEAPLFSOLAKTLFESVQTKAVNKQLDENISLFIHLSPYFLKPK 109

Db 164 sgpckllaladelidsceptqvkummqvlepaqfgdrifslpkelalyvislfpckdlilqa 223

QY 110 AOKCLEWLI 118
 Db 224 aqterywri 232

RESULT 45
 AAY22461
 ID AAY22461 standard; Protein, 627 AA.

XX

RESULT 39
 AAB59198
 ID AAB59198 standard; protein; 589 AA.
 XX
 AC AAB59198;
 XX
 DT 23-MAR-2001 (first entry)
 XX
 DE Human mammary sel-10-1 protein.
 XX
 KW Sel-10; human; Alzheimer's disease; Abeta.
 XX
 OS Homo sapiens.
 XX
 FN W02000754.48-A1
 XX
 FD 14-SEP-2000.
 XX
 DE 23-MAY-2000; 2000W0-US09814
 XX
 PR 09-JUN-1999; 99US-0328877.
 XX
 PA (PHAA) PHARMACIA & UPJOHN CO.
 XX
 PI Pauley AM, Gurney ME, Li J;
 XX
 DT WPT; 2001-102404/11.
 XX
 XX New human sel-10 polypeptides and their encoding polynucleotides,
 PT useful for raising antibodies for detecting sel-10 polypeptide
 PT expression and as drug targets in the treatment of Alzheimer's disease
 PT
 PS Claim 1; Page 79-82; 116pp; English.
 XX
 CC The present invention relates to human sel-10. The sel-10 proteins of
 CC the invention are useful for raising monoclonal or polyclonal
 CC antibodies useful in diagnostic assays for detecting sel-10
 CC polypeptide expression. The sel-10 polypeptides are also useful as drug
 CC targets for decreasing antibody levels in the treatment of Alzheimer's
 CC disease. It is also useful for identifying agents capable of
 CC altering the production level of Abeta. The polynucleotides are useful
 CC for developing assays for identifying agents capable of interfering
 CC with the biological pathways that lead to Alzheimer's disease.
 XX
 SQ Sequence 589 AA;
 Query Match 0.6%; Score 64; DP 22; Length 589;
 Best Local Similarity 27.5%; Pred. No. 86;
 Matches 19; Conservative 10; Mismatches 40; Indels 0; Gaps 0;
 QY 50 TGLELLGLTDPSEFOEAPLESQAKTLERSVQTKAVRKQLDENISLFIHLSPYFLIKP 109
 DB 127 sqpeklalidelidscseptqvlmmqviepfqrdifilpkelalalyvlsflepklbqa 186
 QY 110 AOKCLEWLI 118
 DB 187 aqtorywri 195
 RESULT 40
 AAY22462
 ID AAY22462 standard; protein; 592 AA.
 XX
 AC AAY22462;
 XX
 DT 29-SEP-1999 (first entry)
 XX
 DE Human hippocampal sel-10 protein sequence.
 XX
 KW Sel-10; human; presenilin, Alzheimer's disease; PS-1; PS-2; hippocampus;
 KW therapy.

RESULT 38
 AAB01204
 ID AAB01204 standard; protein; 589 AA.
 XX
 AC AAB01204;
 XX
 DT 03-NOV-2000 (first entry)
 XX
 DE Human GTPase associated protein-29.
 XX
 OS Homo sapiens.
 XX
 FN W0200031263 A2.
 XX
 DT 02-JUN-2000.
 XX
 DE 23-NOV-1999; 99WO-US28013.
 XX
 PR 23 NOV 1998; 98US-0105552.
 XX
 PR 04-FEB-1999; 99US-0118610.
 XX
 PR 06-APR-1999; 99US-0127990
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Hillman JL, Tang YT, Bandman O, Lal P, Yue H, Lu DAM, Baughn MR;
 PT Yang J, Azimzai Y;
 XX
 DT WPT; 2000-400073/34.
 XX
 N-PSDB; AAA49159.
 XX
 CC Human GTPase associated proteins, polynucleotides, and antibodies,
 PT useful for diagnosing, preventing and treating various diseases such as
 PT atherosclerosis, cancer, acquired immune deficiency syndrome (AIDS),
 PT asthma, and autoimmune diseases -
 XX
 PS Claim 1, Page 118 120, 144pp, English.
 XX
 CC Human cDNA libraries from various tissues were screened for GTPase
 CC associated proteins (GTPAP). The present sequence is human
 CC GTPAP-29 protein. This sequence was derived from a cDNA library of
 CC brain tumor tissue. This protein is expressed in nervous,
 CC reproductive and gastrointestinal tissue. The GTPAP proteins may be
 CC used to define agonists and antagonists of GTPAP activity and to
 CC generate antibodies to GTPAP. This means the GTPAP proteins may be
 CC useful for treatment or prevention of diseases associated with GTPAP
 CC such as cell proliferation disorders, autoimmune disorders,
 CC inflammatory disorders, immune system disorders, cancer, AIDS, asthma,
 CC atherosclerosis, arthritis, systemic lupus erythematosus and psoriasis.
 XX
 SQ Sequence 589 AA;
 Query Match 0.6%; Score 64; DB 21; Length 589;
 Best Local Similarity 27.5%; Pred. No. 86;
 Matches 19; Conservative 10; Mismatches 40; Indels 0; Gaps 0;
 QY 50 TGLELLGLTDPSEFOEAPLESQAKTLERSVQTKAVRKQLDENISLFIHLSPYFLIKP 109
 DB 127 sqpeklalidelidscseptqvlmmqviepfqrdifilpkelalalyvlsflepklbqa 186
 QY 110 AOKCLEWLI 118
 DB 187 aqtorywri 195

1

PR 18-JUN-1999; 990S-0139455;
 PR 18-JUN-1999; 990S-0139456;
 PR 18-JUN-1999; 990S-0139457;
 PR 18-JUN-1999; 990S-0139458;
 PR 18-JUN-1999; 990S-0139459;
 PR 18-JUN-1999; 990S-0139460;
 PR 18-JUN-1999; 990S-0139461;
 PR 18-JUN-1999; 990S-0139462;
 PR 18-JUN-1999; 990S-0139463;
 PR 18-JUN-1999; 990S-0139750;
 PR 18-JUN-1999; 990S-0139763;
 PR 21-JUN-1999; 990S-0139817;
 PR 22-JUN-1999; 990S-0139866;
 PR 23-JUN-1999; 990S-0140353;
 PR 23-JUN-1999; 990S-0140354;
 PR 24-JUN-1999; 990S-0140695;
 PR 28-JUN-1999; 990S-0140823;
 PR 28-JUN-1999; 990S-0140961;
 PR 30-JUN-1999; 990S-0141287;
 PR 01-JUL-1999; 990S-0141842;
 PR 01-JUL-1999; 990S-0142154;
 PR 02-JUL-1999; 990S-0142055;
 PR 06-JUL-1999; 990S-0142390;
 PR 08-JUL-1999; 990S-0142802;
 PR 09-JUL-1999; 990S-0142920;
 PR 12-JUL-1999; 990S-0142977;
 PR 13-JUL-1999; 990S-0143452;
 PR 14-JUL-1999; 990S-0143624;
 PR 15-JUL-1999; 990S-0144005;
 PR 16-JUL-1999; 990S-0144085;
 PR 16-JUL-1999; 990S-0144086;
 PR 16-JUL-1999; 990S-0144325;
 PR 19-JUL-1999; 990S-0144331;
 PR 19-JUL-1999; 990S-0144332;
 PR 19-JUL-1999; 990S-0144333;
 PR 19-JUL-1999; 990S-0144334;
 PR 19-JUL-1999; 990S-0144335;
 PR 20-JUL-1999; 990S-0144352;
 PR 20-JUL-1999; 990S-0144632;
 PR 20-JUL-1999; 990S-0144884;
 PR 21-JUL-1999; 990S-0144814;
 PR 21-JUL-1999; 990S-0145086;
 PR 21-JUL-1999; 990S-0145088;
 PR 22-JUL-1999; 990S-0145085;
 PR 22-JUL-1999; 990S-0145087;
 PR 22-JUL-1999; 990S-0145089;
 PR 22-JUL-1999; 990S-0145192;
 PR 23-JUL-1999; 990S-0145145;
 PR 23-JUL-1999; 990S-0145218;
 PR 23-JUL-1999; 990S-0145224;
 PR 26-JUL-1999; 990S-0145276;
 PR 27-JUL-1999; 990S-0145913;
 PR 27-JUL-1999; 990S-0145918;
 PR 27-JUL-1999; 990S-0145919;
 PR 28-JUL-1999; 990S-0145941;
 PR 02-AUG-1999; 990S-0146386;
 PR 02-AUG-1999; 990S-0146388;
 PR 02-AUG-1999; 990S-0146389;
 PR 03-AUG-1999; 990S-0147038;
 PR 04-AUG-1999; 990S-0147204;
 PR 04-AUG-1999; 990S-0147402;
 PR 05-AUG-1999; 990S-0147192;
 PR 05-AUG-1999; 990S-0147260;
 PR 06-AUG-1999; 990S-0147304;
 PR 06-AUG-1999; 990S-0147416;
 PR 09-AUG-1999; 990S-0147A93;
 PR 09-AUG-1999; 990S-0147935;
 PR 10-AUG-1999; 990S-0148171;
 PR 11-AUG-1999; 990S-0148319;
 PR 12-AUG-1999; 990S-0148341;
 PR 13-AUG-1999; 990S-0148565;
 PR 13-AUG-1999; 990S-0148684;
 PR 16-AUG-1999; 990S-0149368;

PR 17-AUG-1999; 990S-0149175;
 PR 18-AUG-1999; 990S-0149426;
 PR 20-AUG-1999; 990S-0149722;
 PR 20-AUG-1999; 990S-0149723;
 PR 20-AUG-1999; 990S-0149929;
 PR 23-AUG-1999; 990S-0149902;
 PR 23-AUG-1999; 990S-0149930;
 PR 25-AUG-1999; 990S-0150566;
 PR 26-AUG-1999; 990S-0150884;
 PR 27-AUG-1999; 990S-0151065;
 PR 27-AUG-1999; 990S-0151066;
 PR 27-AUG-1999; 990S-0151080;
 PR 30-AUG-1999; 990S-0151303;
 PR 31-AUG-1999; 990S-0151438;
 PR 01-SEP-1999; 990S-0151930;
 PR 07-SEP-1999; 990S-0152363;
 PR 10-SEP-1999; 990S-0153070;
 PR 13-SEP-1999; 990S-0153758;
 PR 15-SEP-1999; 990S-0154018;
 PR 16-SEP-1999; 990S-0154039;
 PR 20-SEP-1999; 990S-0154774;
 PR 22-SEP-1999; 990S-0155139;
 PR 23-SEP-1999; 990S-0155486;
 PR 24-SEP-1999; 990S-0155659;
 PR 28-SEP-1999; 990S-0156458;
 PR 29-SEP-1999; 990S-0156596;
 PR 04-OCT-1999; 990S-0157117;
 PR 05-OCT-1999; 990S-0157753;
 PR 06-OCT-1999; 990S-0157865;
 PR 07-OCT-1999; 990S-0158029;
 PR 08-OCT-1999; 990S-0158232;
 PR 08-OCT-1999; 990S-0158369;
 PR 13-OCT-1999; 990S-0159293;
 PR 13-OCT-1999; 990S-0159294;
 PR 13-OCT-1999; 990S-0159329;
 PR 14-OCT-1999; 990S-0159330;
 PR 14-OCT-1999; 990S-0159331;
 PR 14-OCT-1999; 990S-0159637;
 PR 14-OCT-1999; 990S-0159638;
 PR 14-OCT-1999; 990S-0159584;
 PR 21-OCT-1999; 990S-0160741;
 PR 21-OCT-1999; 990S-0160767;
 PR 21-OCT-1999; 990S-0160768;
 PR 21-OCT-1999; 990S-0160814;
 PR 21-OCT-1999; 990S-0160815;
 PR 22-OCT-1999; 990S-0160980;
 PR 22-OCT-1999; 990S-0160981;
 PR 25-OCT-1999; 990S-0161404;
 PR 25-OCT-1999; 990S-0161405;
 PR 25-OCT-1999; 990S-0161406;
 PR 26-OCT-1999; 990S-0161359;
 PR 26-OCT-1999; 990S-0161360;
 PR 26-OCT-1999; 990S-0161361;
 PR 28 OCT-1999; 990S-0161320;
 PR 28 OCT-1999; 990S-0161992;
 PR 28-OCT-1999; 990S-0161993;
 PR 29-OCT-1999; 990S-0162142;

Query Match 0.6%; Score 64; DR 21; Length 109;
 Best Local Similarity 37.5%; Fred. NG. 7.4;
 Matches 15; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

CY 1456 VQHGTSLMNIIVYLLIKIPREKPTIPKAVSNKRSOE 1495
 || | | | | | | | | | | | | | | | | | | | |
 Db 18 vqldqalvavkicgllcdqakktapsvltinqeeqee 57

RESULT 14
 AAG60758
 ID AAG60758 standard; Protein; 109 AA.

PR 05-AUG-1999; 990S-0147152.
 PR 05-AUG-1999; 990S-0147260.
 PR 06-AUG-1999; 990S-0147303.
 PR 06-AUG-1999; 990S-0147416.
 PR 09-AUG-1999; 990S-0147493.
 PR 09-AUG-1999; 990S-0147935.
 PR 10-AUG-1999; 990S-0148171.
 PR 11-AUG-1999; 990S-0148319.
 PR 12-AUG-1999; 990S-0148341.
 PR 13-AUG-1999; 990S-0148565.
 PR 13-AUG-1999; 990S-0148684.
 PR 16-AUG-1999; 990S-0149368.
 PR 17-AUG-1999; 990S-0149375.
 PR 18-AUG-1999; 990S-0149426.
 PR 20-AUG-1999; 990S-0149722.
 PR 20-AUG-1999; 990S-0149723.
 PR 20-AUG-1999; 990S-0149929.
 PR 23-AUG-1999; 990S-0149902.
 PR 23-AUG-1999; 990S-0149930.
 PR 25-AUG-1999; 990S-0150566.
 PR 26-AUG-1999; 990S-0150884.
 PR 27-AUG-1999; 990S-0151065.
 PR 27-AUG-1999; 990S-0151066.
 PR 27-AUG-1999; 990S-0151080.
 PR 30-AUG-1999; 990S-0151303.
 PR 31-AUG-1999; 990S-0151438.
 PR 01-SEP-1999; 990S-0151930.
 PR 07-SEP-1999; 990S-0152363.
 PR 10-SEP-1999; 990S-0153070.
 PR 13-SEP-1999; 990S-0153758.
 PR 15-SEP-1999; 990S-0154018.
 PR 16-SEP-1999; 990S-0154039.
 PR 20-SEP-1999; 990S-0154779.
 PR 22-SEP-1999; 990S-0155139.
 PR 24-SEP-1999; 990S-0155485.
 PR 28-SEP-1999; 990S-0155659.
 PR 29-SEP-1999; 990S-0156458.
 PR 04-OCT-1999; 990S-0157117.
 PR 05-OCT-1999; 990S-0157753.
 PR 06-OCT-1999; 990S-0157865.
 PR 07-OCT-1999; 990S-0158029.
 PR 08-OCT-1999; 990S-0158232.
 PR 12-OCT-1999; 990S-0158369.
 PR 13-OCT-1999; 990S-0159293.
 PR 13-OCT-1999; 990S-0159294.
 PR 13-OCT-1999; 990S-0159295.
 PR 14-OCT-1999; 990S-0159329.
 PR 14-OCT-1999; 990S-0159330.
 PR 14-OCT-1999; 990S-0159331.
 PR 14-OCT-1999; 990S-0159637.
 PR 18-OCT-1999; 990S-0159638.
 PR 21-OCT-1999; 990S-0159584.
 PR 21-OCT-1999; 990S-0160741.
 PR 21-OCT-1999; 990S-0160767.
 PR 21-OCT-1999; 990S-0160768.
 PR 21-OCT-1999; 990S-0160770.
 PR 21-OCT-1999; 990S-0160814.
 PR 21-OCT-1999; 990S-0163817.
 PR 22-OCT-1999; 990S-0165580.
 PR 22-OCT-1999; 990S-0165981.
 PR 22-OCT-1999; 990S-0165989.
 PR 25-OCT-1999; 990S-0161404.
 PR 25-OCT-1999; 990S-0161405.
 PR 25-OCT-1999; 990S-0161406.
 PR 26-OCT-1999; 990S-0161359.
 PR 26-OCT-1999; 990S-0161360.
 PR 26-OCT-1999; 990S-0161361.
 PR 28-OCT-1999; 990S-0161920.
 PR 28-OCT-1999; 990S-0161992.
 PR 28-OCT-1999; 990S-0161993.
 PR 29-OCT-1999; 990S-0162142.

Query Match 0.68, Score 66, DB 21, Length 250;
 Best Local Similarity 19.3%; Pred. No. 13;
 Matches 16, Conservative 25, Mismatches 42, Indels 0, Gaps 0;
 QY 1828 KTYKQTEKKNHMGPFMSLQEHGKMKKEELTSHQSQTATFFALDFRAQISENLE 1887
 DB 33 reylklkgrnyenlgrqqrlljedbfinskeleqlerqlqdsikpryrcktyymtqls 152

QY 1888 EVGKTENCILICIVAMVVKLSEV 1910
 DB 153 dlqgkehilldanralsmkledm 175

RESULT 11
 AAY84910
 ID AAY84910 standard; Protein: 250 AA.
 XX
 AC AAY84910;
 XX
 PT 21-AUG-2000 (first entry)
 XX
 DE Amino acid sequence of the Arabidopsis AGL4 protein.
 XX
 KW AGL4; promoter; transgenic plant; suppressed flowering; wood;
 KW agamous-like protein, floral organ selective regulatory element.
 XX
 OS Arabidopsis thaliana.
 XX
 PN W026002:578-A2.
 XX
 PD 27-APR-2000.
 XX
 PF 15 OCT-1999; 990S-01524407
 XX
 PP 16-OCT-1999; 990S-0104604.
 XX
 PA (PESC) UNIV CALIFORNIA.
 XX
 PI Yanofsky MF;
 XX
 PR WPI: 2000-330680/29.
 PR N-PSDR: AAA15013.
 XX
 PT Transgenic plants in which flowering is suppressed by a tissue specific
 PT cytotoxic gene product, useful for the production of wood for use as
 PT lumber or pulp.
 XX
 PS Disclosure: Fig 10a-b; 79pp; English.
 XX
 CC The present sequence represents the Arabidopsis agamous like (AGL) 4
 CC protein. The AGL4 promoter is used to produce transgenic plants of the
 CC invention, which are characterized by suppressed flowering. The plants
 CC are transformant with a construct comprising a floral organ selective
 CC regulatory element (e.g. AGL4 promoter), operatively linked to a
 CC nucleotide sequence encoding a cytotoxic gene product (which is
 CC inheritable by the progeny). The plants may be grown for either human
 CC consumption or for use as a raw material in industry. When trees, they
 CC are particularly suitable for cultivation to provide wood. As the
 CC flowering process consumes 20-30% of the energy of a typical plant,
 CC it is advantageous to suppress flowering in order to improve wood and
 CC lumber yields. Suppression of flowering may be desired to eliminate
 CC the production of allergic pollen, or to prevent pollen dissemination.
 XX
 SQ Sequence 250 AA.

Query Match 0.68, Score 66, DB 21, Length 250;
 Best Local Similarity 19.3%; Pred. No. 13;
 Matches 16, Conservative 25, Mismatches 42, Indels 0, Gaps 0;
 QY 1828 KTYKQTEKKNHMGPFMSLQEHGKMKKEELTSHQSQTATFFALDFRAQISENLE 1887
 DB 93 reylklkgrnyenlgrqqrlljedbfinskeleqlerqlqdsikpryrcktyymtqls 152

PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138543.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139392.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.
PR 18-JUN-1999; 9905-0139817.
PR 21-JUN-1999; 9905-0139899.
PR 22-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 23-JUN-1999; 9905-0140355.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0143977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143621.
PR 15-JUL-1999; 9905-0144055.
PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144884.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145385.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145218.
PR 23-JUL-1999; 9905-0145224.
PR 26-JUL-1999; 9905-0145276.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145918.
PR 28-JUL-1999; 9905-0145919.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146388.
PR 02-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147204.
PR 05-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 06-AUG-1999; 9905-0147493.
PR 09-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.
PR 10-AUG-1999; 9905-0148319.
PR 11-AUG-1999; 9905-0148341.
PR 12-AUG-1999; 9905-0148565.
PR 13-AUG-1999; 9905-0148684.
PR 13-AUG-1999; 9905-0149368.
PR 16-AUG-1999; 9905-0149175.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 20-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 01-SEP-1999; 9905-0152363.
PR 07-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157865.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 22-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161992.


```

XX OS Neisseria gonorrhoeae.
XX PN WC9924578-A2.
XX XX
XX PD 20-MAY-1999.
XX XX
XX PF 09-OCT-1998; 98WO-1801665.
XX PR 01-SEP-1998; 98GB-0019016.
XX PR 06-NOV-1997; 97GB-0023516.
XX PR 14-NOV-1997; 97GB-0024190.
XX PR 18-NOV-1997; 97GB-0024386.
XX PR 27-NOV-1997; 97GB-0025158.
XX PR 10-DEC-1997; 97GB-0026147.
XX PR 14-JAN-1998; 98GB-0000759.
XX PA (CHIR-) CHIRON SPA.
XX XX
XX PI Grandi G, Masiagnani V, Pizza M, Rappuoli R, Scarlato V;
XX XX
XX PF WPI; 1999-327407/27.
XX PR P-PSDB; AAY38789.
XX XX
XX PT Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
XX PT diagnosis, treatment and prevention of infection
XX XX
XX PS Claim 4; Page 332; 524pp; English.
XX XX
XX CC Amino acid sequences AAY38499-Y38944 represent Neisseria meningitidis
XX CC and N. gonorrhoeae antigenic proteins. They are encoded by open
XX CC reading frames (ORFs) AAC11972-212358. The antigenic proteins,
XX CC their fragments, their nucleic acids and antibodies are used for
XX CC diagnosis, prevention (as vaccines) or treatment of Neisseria
XX CC infections, such as meningitis, septicemia and gonorrhea. Both
XX CC organisms are closely related. Fragments of the nucleic acids
XX CC are useful as hybridisation probes and antisense reagents.
XX XX
XX SQ Sequence 513 AA;

Query Match 0.69; Score 67; DB 20; Length 513;
Best local Similarity 44.79; Pred. No. 27;
Matches 17; Conservative 5; Mismatches 16; Indels 0; Gaps 0;

QY 184 KDLGPMDFICSLVTKSVKVFPAEYFSSAQLEVLAFYA 221
: | | | | | | | | | | | | | | | |
Db 349 rglvflpfmvsprvsagvlllypgwtasipillamy 386

RESULT 6
AAY38789
ID AAY38789 standard; Protein: 513 AA.
XX AC AAY38789;
XX DT 08-OCT-1999 (first entry)
XX DE Neisseria gonorrhoeae antigenic protein encoded by ORF139.
XX KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
XX KW treatment; Neisseria infection; meningitis; septicemia; gonorrhea.
XX OS Neisseria gonorrhoeae.
XX PN WC9924578-A2.
XX PD 20-MAY-1999.
XX PF 09-OCT-1998; 98WO-1801665.
XX PR 01-SEP-1998; 98GB-0019016.
XX PR 06-NOV-1997; 97GB-0023516.

```

```

PR 14-NOV-1997; 97GB-0024190.
PR 18-NOV-1997; 97GB-0024386.
PR 27-NOV-1997; 97GB-0025158.
PR 10-DEC-1997; 97GB-0026147.
PR 14-JAN-1998; 98GB-0000759.
XX PA (CHIR-) CHIRON SPA.
XX XX
XX PI Grandi G, Masiagnani V, Pizza M, Rappuoli R, Scarlato V;
XX XX
XX PF WPI; 1999-327407/27.
XX PR P-PSDB; AAY38789.
XX XX
XX PT Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
XX PT diagnosis, treatment and prevention of infection
XX XX
XX PS Claim 4; Page 332; 524pp; English.
XX XX
XX CC Amino acid sequences AAY38499-Y38944 represent Neisseria meningitidis
XX CC and N. gonorrhoeae antigenic proteins. They are encoded by open
XX CC reading frames (ORFs) AAC11972-212358. The antigenic proteins,
XX CC their fragments, their nucleic acids and antibodies are used for
XX CC diagnosis, prevention (as vaccines) or treatment of Neisseria
XX CC infections, such as meningitis, septicemia and gonorrhea. Both
XX CC organisms are closely related. Fragments of the nucleic acids
XX CC are useful as hybridisation probes and antisense reagents.
XX XX
XX SQ Sequence 513 AA;

Query Match 0.69; Score 67; DB 20; Length 513;
Best local Similarity 44.79; Pred. No. 27;
Matches 17; Conservative 5; Mismatches 16; Indels 0; Gaps 0;

QY 184 KDLGPMDFICSLVTKSVKVFPAEYFSSAQLEVLAFYA 221
: | | | | | | | | | | | | | | | |
Db 349 rglvflpfmvsprvsagvlllypgwtasipillamy 386

RESULT 7
AAY29083
ID AAY29083 standard; Protein: 1199 AA.
XX AC AAY29083;
XX DT 27-SEP-1999 (first entry)
XX DE Human C1F150/hTAF1150 protein.
XX KW C1F150/hTAF1150; mitosis; cell cycle progression; neoplasia; human;
XX KW transcriptional regulator; benign proliferative disease; dysplastic;
XX KW hyperplastic disorder; anemia.
XX OS Homo sapiens.
XX PN WO9933985-A2.
XX PD 08-JUL-1999.
XX PF 28-DEC-1998; 98WO-0527665.
XX PR 30-NOV-1997; 97NS-0068956.
XX PA (CHIR-) CHIRON COPP.
XX PI Kaufmann J;
XX DR WPI; 1999-444069/37.
XX DR N-PSDB; AAX91483.
XX PT New isolated ICIF150/hTAF1150 polypeptides
XX PS Claim 2; Page 78; 83pp; English.

```


XX Baer R, Bowcock AM;
XX WPI: 1998-230417/20;
XX N-PSDB; AAV24135;
XX
XX DNA sequence encoding BARD1, B123, BE2, BE14, BE31 or BE445 - which
XX as breast cancer antigen, BRCAL, binding proteins are useful to
XX identify patient having or at risk of developing cancer
XX
XX Discl:see Pat. 287,282, 348ff, English.
XX
XX The sequence is that of a protein which can be used in the
XX preparation of the recombinant breast cancer antigen, BRCAL, binding
XX proteins BARD1, B123, BE2, BE14, BE31 or BE445, of a composition for the
XX detection of a BARD1, B123, BE2, BE14, BE31 or BE445 nucleic acid
XX sequence, specifically a wild type BARD1 composition for the detection
XX or purification of BRCAL, useful to identify a patient having, or at
XX risk of developing cancer. BARD1 can be used in the preparation of an
XX anti-BARD1 antibody, and in the detection and purification of an
XX protein. BARD1, B123, BE2, BE14, BE31 or BE445 can be used in the
XX identification of a binding protein agonist or antagonist that alters
XX the binding of BARD1, B123, BE2, BE14, BE31 or BE445 to BRCAL or the
XX biological activity of the BRCAL-BARD1, B123, BE2, BE14, BE31 or BE445
XX complex. The antibodies can be used to detect BARD1, B123, BE2, BE14,
XX BE31 or BE445, a specific anti-BARD1 antibody can be used to identify
XX a patient having or at risk of developing cancer.
XX
SQ Sequence 515 AA:

Query Match 23.9%; Score 2585; DB 14; Length 515;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1540 RLSKPEIVITPEFLPVSLIAVPEPFPEDEPEPAIPEQALVYIFKIDKSPGAPNPEPF 1649
DB 1015KPEIVITPEFLPVSLIAVPEPFPEDEPEPAIPEQALVYIFKIDKSPGAPNPEPF 60
QY 1640 VIVIXTAVKLTAPKPKPKNVIGSALICIAVFTSTLPAALIPQIPSIPTTKNSP 1749
DB 61 VIVIXTAVKLTAPKPKPKNVIGSALICIAVFTSTLPAALIPQIPSIPTTKNSP 120
QY 1750 IVSSKPVYLLSALAAIKQVETIPPEISPTVEGIIISQVTHIEKITSEM;SASQANIPKLSI 1809
DB 121 IVSSKPVYLLSALAAIKQVETIPPEISPTVEGIIISQVTHIEKITSEM;SASQANIPKLSI 180
QY 1810 KETIATTAAPPVLLPAITPTTYKQITPKNPNHNSPPMSITQPHITXMKKEEITSHSGLTA 1869
DB 181 KETIATTAAPPVLLPAITPTTYKQITPKNPNHNSPPMSITQPHITXMKKEEITSHSGLTA 240
QY 1870 FFLRALDEKAQHSNDEGFAVCKIRNIIILAGLVAMVVKISEVIERPEFEKFLPWAKIEDAP 1929
DB 241 FFLRALDEKAQHSNDEGFAVCKIRNIIILAGLVAMVVKISEVIERPEFEKFLPWAKIEDAP 300
QY 1930 KPEKIDFYNAIPALTAERLKGEPFLFAGHLVKEPAPTLXQVNIISFTIEAPFSENPKECC 1989
DB 301 KPEKIDFYNAIPALTAERLKGEPFLFAGHLVKEPAPTLXQVNIISFTIEAPFSENPKECC 360
QY 1990 LLIQELNPTTPYFLPTGHPISPPSAXALMMPIVQILNPLGCPKPKTGPVTRHLIPCT 2049
DB 361 LLIQELNPTTPYFLPTGHPISPPSAXALMMPIVQILNPLGCPKPKTGPVTRHLIPCT 420
QY 2070 AGPSVANAGQELMRFVNAQGLLEFDESSPKVSEFVALLVIALAEKLNIVLLSESPK 2199
DB 421 AGPSVANAGQELMRFVNAQGLLEFDESSPKVSEFVALLVIALAEKLNIVLLSESPK 480
QY 2110 LAELMEDETEVEHCQKQITQGLTEVTLGELQSYT 2144
DB 481 LAELMEDETEVEHCQKQITQGLTEVTLGELQSYT 515

RESULT 2

AAB43598
ID AAB43598 standard; Protein; 343 AA.
XX
XX AAB43598;
XX
XX 08-FEB-2001 (first entry)
XX
XX Human cancer associated protein sequence SW2 ID B01043.
DE
XX
XX Human; cancer associated gene, cancer antigen; detection; cancer;
XX diagnosis; metastatic; proliferative; vulvovaginary; immunomodulator;
XX antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
XX antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
XX dermatological; neuroprotective; thrombolytic; coagulant; neotropic;
XX vasodilator; antiproliferative; antineoplastic; gene therapy; inflammation;
XX immune disorder; haematopoietic cell disorder; autoimmune disorder;
XX allergic reaction; graft versus host disease; organ rejection;
XX haemostatic; thrombolytic; cardiovascular disorder; infection;
XX neurological disease; drug screening.
XX
XX Homo sapiens
XX
XX WC2000055350-AL.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US05882.
XX
XX 12-MAR-1999; 900G-01.4770.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI: 2000-587533/55.
XX
XX N-PSDB; AA:77807.
XX
XX Novel isolated nucleic acid comprising sequences encoding peptides
XX useful for treating or diagnosing e.g. cancer -
XX
XX Claim 17; Page 1631-1632; 2352pp; English.
XX
XX AAC77607 to AAC78448 encode the human cancer associated proteins given
XX in AAB43398 to AAB44239. The proteins can have activities based on the
XX tissues and cells the genes are expressed in. Example of activities
XX include cytostatic; proliferative; vulvovaginary; immunomodulator;
XX antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
XX antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
XX dermatological; neuroprotective; cardiac; thrombolytic; coagulant;
XX vasodilator; vasotropic; antiproliferative and antineoplastic. The
XX polypeptides and polypeptides can be used for preventing, treating or
XX ameliorating medical conditions and diagnosing pathological conditions.
XX Polynucleotides, polypeptides, antibodies, agonists and antagonists from
XX the present invention may be used to treat immune disorders by activating
XX or inhibiting the proliferation, differentiation or mobilisation of
XX immune cells, to treat disorders of haematopoietic cells, autoimmune
XX disorders, allergic reactions, graft versus host disease and organ
XX rejection, modulate haemostatic or thrombolytic activity, modulate
XX inflammation, cancers, cardiovascular disorders, neurological disease and
XX bacterial or viral infections. The peptides, nucleotides, antibodies,
XX agonists and antagonists may be also be used in drug screens. AAC78449 to
XX AAC78457 and AAB44240 represent sequences used in the identification of
XX the present invention.
XX
XX Sequence 343 AA:

Query Match 0.7%; Score 71; DB 21; Length 343;
Best Local Similarity 37.0%; Pred. No. 4.3;
Matches 17; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

QY 1813 LATTIAPRVLLPAIKTKYKQIEKNKNHNGPMSILOEHIGXMKKE 1858
DB 1813 LATTIAPRVLLPAIKTKYKQIEKNKNHNGPMSILOEHIGXMKKE 1858

65	60	0.5	427	12	AAV14346	Sequence of the po	158	0.5	613	14	AAV14347	CKS-HCV and upon fu
66	60	0.5	427	17	AAV6326	P. dentalis haem	159	0.5	613	22	AAV6326	HCV recombinant an
67	60	0.5	427	19	AAV6463	Buematulintin pro	160	0.5	781	13	AAV21665	HCV CKS-48-10-1 p
68	60	0.5	427	20	AAV6876	Nosseria meningit	161	0.5	781	13	AAV22208	Sequence of fusion
69	60	0.5	427	21	AAV6876	Amino acid sequen	162	0.5	781	14	AAV33632	HCV CKS-48-10-1 p
70	60	0.5	427	22	AAV6876	Sequence of proteo	163	0.5	781	14	AAV33632	HCV CKS-48-10-1 p
71	60	0.5	427	23	AAV6876	Human foetal cell	164	0.5	781	14	AAV33632	HCV CKS-48-10-1 p
72	60	0.5	427	24	AAV6876	Chicken derived ne	165	0.5	781	15	AAV33632	HCV CKS-48-10-1 p
73	60	0.5	427	25	AAV6876	Retinal cancer asso	166	0.5	781	22	AAV33632	HCV recombinant an
74	60	0.5	427	26	AAV6876	Merge protein tyro	167	0.5	859	10	AAV90146	Seq. extending thro
75	60	0.5	427	27	AAV6876	Murine protein tyro	168	0.5	859	10	AAV90146	Sequence encoded i
76	60	0.5	427	28	AAV6876	Mouse related adhe	169	0.5	882	15	AAV45749	Alpha-DNA polymera
77	60	0.5	427	29	AAV6876	Human foetal cell	170	0.5	971	14	AAV33632	CKS-HCV and upon fu
78	60	0.5	427	30	AAV6876	P. dentalis haem	171	0.5	971	22	AAV6326	HCV recombinant an
79	60	0.5	427	31	AAV6876	Buematulintin pro	172	0.5	971	22	AAV6326	CKS-HCV and upon fu
80	60	0.5	427	32	AAV6876	Sequence of hepari	173	0.5	971	22	AAV6326	HCV recombinant an
81	60	0.5	427	33	AAV6876	Sequence of the po	174	0.5	992	22	AAV6326	CKS-HCV and upon fu
82	60	0.5	427	34	AAV6876	Polyptide p1689	175	0.5	992	22	AAV6326	HCV recombinant an
83	60	0.5	427	35	AAV6876	P1689 HCV antipe	176	0.5	1021	18	AAV44481	Fusion protein c28
84	60	0.5	427	36	AAV6876	Sequence of hepari	177	0.5	1021	19	AAV44481	Sequence encoded i
85	60	0.5	427	37	AAV6876	Sequence of the po	178	0.5	1766	16	AAV20241	Tuberculous serosis
86	60	0.5	427	38	AAV6876	Arabidopsis thalia	179	0.5	1766	16	AAV20241	H. pylori GHP-1
87	60	0.5	427	39	AAV6876	Arabidopsis thalia	180	0.5	1766	16	AAV20241	True type A polio
88	60	0.5	427	40	AAV6876	Plant SEF encoded	181	0.5	2261	13	AAV22210	Peptide encoded by
89	60	0.5	427	41	AAV6876	Plant SEF encoded	182	0.5	2261	13	AAV22210	Sequence encoded i
90	60	0.5	427	42	AAV6876	Peptide encoded by	183	0.5	2301	10	AAV22210	HCV polypeptide 1
91	60	0.5	427	43	AAV6876	Human foetal cell	184	0.5	2301	10	AAV22210	Sequence encoded i
92	60	0.5	427	44	AAV6876	Chlamydia trachoma	185	0.5	2301	10	AAV22210	Sequence encoded i
93	60	0.5	427	45	AAV6876	Plant SEF encoded	186	0.5	2301	10	AAV22210	HCV amino acid seq
94	60	0.5	427	46	AAV6876	Plant SEF encoded	187	0.5	2301	10	AAV22210	Peptide encoded by
95	60	0.5	427	47	AAV6876	Plant SEF encoded	188	0.5	2301	10	AAV22210	Sequence encoded i
96	60	0.5	427	48	AAV6876	Arabidopsis thalia	189	0.5	2549	16	AAV1740	Stimulus effect of
97	60	0.5	427	49	AAV6876	Arabidopsis thalia	190	0.5	2549	16	AAV1740	Human RAP11. Homo
98	60	0.5	427	50	AAV6876	Arabidopsis thalia	191	0.5	2549	16	AAV1740	FRAP (for1) protei
99	60	0.5	427	51	AAV6876	Arabidopsis thalia	192	0.5	2549	16	AAV1740	Hepatitis C virus
100	60	0.5	427	52	AAV6876	Open reading frame	193	0.5	2772	21	AAV18540	Protein encoded by
101	60	0.5	427	53	AAV6876	HCV antigen C100	194	0.5	2772	21	AAV18540	HCV-1 polyprotein
102	60	0.5	427	54	AAV6876	C. terminalis of supe	195	0.5	2894	13	AAV24440	Composite HCV HCV
103	60	0.5	427	55	AAV6876	Sequence of the su	196	0.5	2894	13	AAV24440	Composite HCV HCV
104	60	0.5	427	56	AAV6876	Arabidopsis thalia	197	0.5	2955	11	AAV08124	Hepatitis C virus
105	60	0.5	427	57	AAV6876	Streptomyces viola	198	0.5	2955	20	AAV14975	Amino acid sequen
106	60	0.5	427	58	AAV6876	Ammonia acid sequen	199	0.5	2955	21	AAV18541	Polyptide encoded
107	60	0.5	427	59	AAV6876	Arabidopsis thalia	200	0.5	4011	14	AAV21519	Compiled HCV sequen
108	60	0.5	427	60	AAV6876	Arabidopsis thalia		0.5	4011	14	AAV21519	Hepatitis C virus

ALIGNMENTS

RESULT	1
AAW54099	AAW54099 standard; Protein: 515 AA.
XX	AAW54099
XX	AAW54099
XX	28-SEP-1998 (first entry)
XX	Homo sapiens BAP28 sequence.
XX	Homo sapiens BAP28 sequence.
XX	BAP28; ring protein; BPCAL; Breast cancer; risk; diagnosis.
XX	Homo sapiens.
XX	Homo sapiens.
XX	W9812127 A2.
XX	26-MAR-1998.
XX	19-SEP-1997; 97WO-US16842.
XX	04-APR-1997; 97US-0042985.
XX	20-SEP-1996; 96US-0025296.
XX	03-APR-1997; 97US-0042611.
XX	(TEXA) UNIV TEXAS SYSTEM.

101	58	0.5	4011	4	US-09-014-416-5	Sequence 5, Appl	174	56	0.5	3011	1	US-08-188-281B-1	Sequence 1, Appl
102	58	0.5	3012	3	US-08-811-566-2	Sequence 2, Appl	175	56	0.5	3011	1	US-08-453-552-1	Sequence 1, Appl
103	57	0.5	200	1	US-08-209-182C-8	Sequence 8, Appl	176	56	0.5	3011	2	US-08-710-637-1	Sequence 1, Appl
104	57	0.5	201	5	PCT-US92-05612-8	Sequence 8, Appl	177	56	0.5	3011	5	PCT-US93-00907-1	Sequence 1, Appl
105	57	0.5	231	3	US-08-771-098-2	Sequence 4, Appl	178	56	0.5	3011	5	PCT-US94-07280-1	Sequence 1, Appl
106	57	0.5	231	3	US-09-022-875-4	Sequence 2, Appl	179	56	0.5	3011	5	PCT-US95-01087-1	Sequence 1, Appl
107	57	0.5	231	4	US-09-354-040-2	Sequence 2, Appl	180	56	0.5	3033	1	US-07-925-695-5	Sequence 105, App
108	57	0.5	269	4	US-09-028-366-6	Sequence 6, Appl	181	55	0.5	107	1	US-08-182-175A-105	Sequence 92, App
109	57	0.5	326	4	US-09-058-389A-3	Sequence 3, Appl	182	55	0.5	107	5	US-08-474-643A-92	Sequence 105, App
110	57	0.5	351	1	US-08-402-217A-2	Sequence 2, Appl	183	55	0.5	235	2	PCT-US92-06412-105	Sequence 2, Appl
111	57	0.5	351	1	US-08-700-178-2	Sequence 2, Appl	184	55	0.5	336	2	US-08-474-379C-88	Sequence 88, Appl
112	57	0.5	351	3	US-08-995-654-2	Sequence 2, Appl	185	55	0.5	479	4	US-09-177-349-3	Sequence 4, Appl
113	57	0.5	397	1	US-08-415-823-2	Sequence 2, Appl	186	55	0.5	536	5	PCT-US95-05008-12	Sequence 22, Appl
114	57	0.5	397	2	US-09-086-662-2	Sequence 2, Appl	187	55	0.5	536	5	US-07-688-352C-32	Sequence 85, Appl
115	57	0.5	465	2	US-08-162-402B-8	Sequence 16, Appl	188	55	0.5	734	3	US-09-146-249A-85	Sequence 85, Appl
116	57	0.5	570	2	US-08-484-158B-16	Sequence 16, Appl	189	55	0.5	734	3	US-08-206-188B-85	Sequence 2, Appl
117	57	0.5	570	2	US-08-484-596A-16	Sequence 16, Appl	190	55	0.5	857	1	US-07-717-331F-2	Sequence 65, Appl
118	57	0.5	570	2	US-08-480-150A-16	Sequence 16, Appl	191	55	0.5	886	3	US-08-474-379C-65	Sequence 65, Appl
119	57	0.5	570	3	US-08-458-731-16	Sequence 16, Appl	192	55	0.5	886	3	US-09-146-249A-85	Sequence 65, Appl
120	57	0.5	570	3	US-08-149-224A-16	Sequence 16, Appl	193	55	0.5	886	3	US-08-206-188B-85	Sequence 12, Appl
121	57	0.5	591	1	US-08-145-955A-21	Sequence 21, Appl	194	55	0.5	898	2	US-08-474-379C-12	Sequence 12, Appl
122	57	0.5	591	2	US-08-451-747-21	Sequence 21, Appl	195	55	0.5	898	2	US-09-146-249A-12	Sequence 12, Appl
123	57	0.5	591	3	US-09-134-853-21	Sequence 21, Appl	196	55	0.5	898	3	US-08-206-188B-12	Sequence 12, Appl
124	57	0.5	630	4	US-09-300-909-20	Sequence 20, Appl	197	55	0.5	900	1	US-07-688-352C-12	Sequence 12, Appl
125	57	0.5	684	4	US-08-925-762-2	Sequence 2, Appl	198	55	0.5	942	3	US-08-560-005-7	Sequence 6, Appl
126	57	0.5	770	2	US-08-299-521-13	Sequence 13, Appl	199	55	0.5	968	3	US-08-560-005-7	Sequence 7, Appl
127	57	0.5	770	2	US-08-961-810-123	Sequence 123, App	200	55	0.5				
128	57	0.5	770	4	US-08-352-902B-123	Sequence 123, App							
129	57	0.5	852	2	US-09-070-060-3	Sequence 3, Appl							
130	57	0.5	852	3	US-09-357-746-3	Sequence 3, Appl							
131	57	0.5	852	3	US-08-775-009-34	Sequence 34, Appl							
132	57	0.5	888	1	US-08-445-640-35	Sequence 35, Appl							
133	57	0.5	888	3	US-08-170-558-35	Sequence 35, Appl							
134	57	0.5	888	3	US-08-447-314-35	Sequence 35, Appl							
135	57	0.5	888	3	US-08-445-461-35	Sequence 35, Appl							
136	57	0.5	1240	3	US-08-530-956A-4	Sequence 4, Appl							
137	57	0.5	2487	2	US-08-970-269A-29	Sequence 29, Appl							
138	57	0.5	2487	4	US-09-167-562-29	Sequence 29, Appl							
139	57	0.5	3033	1	US-07-925-695-9	Sequence 9, Appl							
140	57	0.5	3033	1	US-08-970-269A-30	Sequence 30, Appl							
141	57	0.5	3959	2	US-09-407-562-30	Sequence 30, Appl							
142	57	0.5	3959	4	US-09-036-987A-5	Sequence 5, Appl							
143	57	0.5	4928	4	US-09-370-700-5	Sequence 5, Appl							
144	57	0.5	4928	4	US-09-100-804-20	Sequence 20, Appl							
145	57	0.5	75	3	US-08-444-818-14	Sequence 14, Appl							
146	56	0.5	237	4	US-08-701-191A-21	Sequence 21, Appl							
147	56	0.5	237	2	US-08-989-478-3	Sequence 3, Appl							
148	56	0.5	314	2	US-08-956-685-3	Sequence 3, Appl							
149	56	0.5	314	3	US-08-379-556A-8	Sequence 8, Appl							
150	56	0.5	333	2	US-08-390-000A-8	Sequence 8, Appl							
151	56	0.5	348	2	US-08-430-286A-11	Sequence 11, Appl							
152	56	0.5	406	3	US-09-141-047-9	Sequence 9, Appl							
153	56	0.5	419	1	US-07-925-695-9	Sequence 9, Appl							
154	56	0.5	489	1	US-08-656-034-16	Sequence 16, Appl							
155	56	0.5	548	3	US-08-948-504-16	Sequence 16, Appl							
156	56	0.5	576	3	US-08-937-540-4	Sequence 4, Appl							
157	56	0.5	591	2	US-08-656-034-2	Sequence 2, Appl							
158	56	0.5	624	4	US-08-190-802A-63	Sequence 63, Appl							
159	56	0.5	713	1	US-08-477-346-63	Sequence 6, Appl							
160	56	0.5	713	4	US-09-192-083-6	Sequence 6, Appl							
161	56	0.5	743	4	US-08-257-073-5	Sequence 5, Appl							
162	56	0.5	740	1	US-08-249-380-2	Sequence 2, Appl							
163	56	0.5	968	2	US-08-588-985-2	Sequence 2, Appl							
164	56	0.5	1865	1	US-08-971-588-2	Sequence 4, Appl							
165	56	0.5	1865	2	US-08-619-554-4	Sequence 3, Appl							
166	56	0.5	1895	2	US-08-546-291-3	Sequence 3, Appl							
167	56	0.5	2465	3	US-09-160-804-3	Sequence 2, Appl							
168	56	0.5	2465	3	PCT-US94-0394-3	Sequence 2, Appl							
169	56	0.5	2465	4	US-09-290-640-46	Sequence 46, Appl							
170	56	0.5											
171	56	0.5											
172	56	0.5											
173	56	0.5											

ALIGNMENTS

RESULT 1
 US-08-999-774A-13
 : Sequence 13, Application US/08999774A
 : Patent No. 6274313
 : GENERAL INFORMATION:
 : APPLICANT: Gish, Kurt C.
 : APPLICANT: Schickel, Wolfgang
 : APPLICANT: Shanahan, Frances
 : APPLICANT: Lees, Emma M.
 : APPLICANT: McLanahan, Terrill K.
 : TITLE OF INVENTION: Intracellular Regulatory Molecules;
 : TITLE OF INVENTION: Related Reagents
 : NUMBER OF SEQUENCES: 13
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: DNAX Research Institute
 : STREET: 901 California Avenue
 : CITY: Palo Alto
 : STATE: California
 : COUNTRY: USA
 : ZIP: 94304-1104
 : COMPUTER READABLE FORM: disk
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US-08-999,774A
 : FILING DATE: 10-DEC-1997
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 60/042,818
 : FILING DATE: 11-DEC-1996
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Chiu, Edwin P.
 : REGISTRATION NUMBER: 34,090
 : REFERENCE/WORK KEY NUMBER: DX0646
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (650)852-9196
 : TELEFAX: (650)456-1200
 : INFORMATION FOR SEQ ID NO: 13:
 : SEQUENCE CHARACTERISTICS:

us-09-603-665-5-rai

Fri Nov 9 15:42:38 2001

APPLICATION NUMBER: US 07/647,831
FILING DATE: 29-JAN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H.
REGISTRATION NUMBER: 29,959
REFERENCE/CKET NUMBER: 2415 20022 00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-813-5600
TELEFAX: 415-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 343 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-047-413-13

Query Match 0.6%; Score 64; DB 1; Length 343;
Best Local Similarity 31.0%; Pred. No. 15;
Matches 13; Conservative 8; Mismatches 21; Indels 0; Gaps 0;

QY 1428 VLAAYGCKDALEAVTFWFVSVCFPSVQHQIQSLMNILOY 1469
:||| | | | | : : : : | | | | : | : |
DB 167 ILARTYPRKDSIFRCEPNSFINPIIOFLVQHNLPLLANVYP 208

RESULT 6

US-08-229-050-13
Sequence 13, Application US/08224050
Patent No. 606491
GENERAL INFORMATION:
APPLICANT: Cornelissen, Bernardus J.C.
APPLICANT: Melchers, Leo S.
APPLICANT: Meulenhoff, Elisabeth J.S.
APPLICANT: van Koekel, Jeroen S.C.
APPLICANT: Sela-Buurlage, Marianne B.
APPLICANT: Vloemans, Alexandra A.
APPLICANT: Woloshuk, Charles P.
APPLICANT: Bol, John F.
APPLICANT: Linthorst, Hubertus J.M.
TITLE OF INVENTION: FUNGAL RESISTANT PLANTS, PROCESS FOR
OBTAINING FUNGAL RESISTANT PLANTS AND RECOMBINANT
TITLE OF INVENTION: POLYNUCLEOTIDES FOR USE THEREIN
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESS: Morrison & Foerster
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/229,050
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,831
FILING DATE: 29-JAN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H.
REGISTRATION NUMBER: 29,959
REFERENCE/CKET NUMBER: 2415 20022 00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-813-5600
TELEFAX: 415-494-0792
TELEX: 706141

INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 343 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-229-050-13

Query Match 0.6%; Score 64; DB 3; Length 343;
Best Local Similarity 31.0%; Pred. No. 15;
Matches 13; Conservative 8; Mismatches 21; Indels 0; Gaps 0;

QY 1428 VLAAYGCKDALEAVTFWFVSVCFPSVQHQIQSLMNILOY 1469
:||| | | | | : : : : | | | | : | : |
DB 167 ILARTYPRKDSIFRCEPNSFINPIIOFLVQHNLPLLANVYP 208

RESULT 7

US-08-801-563-13
Sequence 13, Application US/08801563
Patent No. 6087560
GENERAL INFORMATION:
APPLICANT: Cornelissen, Bernardus J.C.
APPLICANT: Melchers, Leo S.
APPLICANT: Meulenhoff, Elisabeth J.S.
APPLICANT: van Koekel, Jeroen S.C.
APPLICANT: Sela-Buurlage, Marianne B.
APPLICANT: Vloemans, Alexandra A.
APPLICANT: Woloshuk, Charles P.
APPLICANT: Bol, John F.
APPLICANT: Linthorst, Hubertus J.M.
TITLE OF INVENTION: FUNGAL RESISTANT PLANTS, PROCESS FOR
OBTAINING FUNGAL RESISTANT PLANTS AND RECOMBINANT
TITLE OF INVENTION: POLYNUCLEOTIDES FOR USE THEREIN
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESS: Morrison & Foerster
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,563
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/047,413
FILING DATE: 19-APR-1993
APPLICATION NUMBER: US 07/647,831
FILING DATE: 29-JAN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H.
REGISTRATION NUMBER: 29,959
REFERENCE/CKET NUMBER: 2415 20022 00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-813-5600
TELEFAX: 415-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 343 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-801-563-13

us-09-603-665-5.ra1

Fri Nov 9 15:42:38 2001

```

; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: 5-Ethylpurylskimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,008
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/306,063
; FILING DATE: 13-SEP-1994
; APPLICANT: US 07/749,611
; FILING DATE: 28-AUG-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(10660)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6099
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 426 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-008-63

Query Match 0.88; Score 60; DB 1; Length 426;
Best Local Similarity 47.8%; Pred. No. 78;
Matches 11; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Q7 2028 ENPIGGEKFGPPVTKH:IPCTA 2050
Db 111 EYMLGGEPMEERPIGHLVDCLA 133

RESULT 18
US-08-406-008-63
; Sequence 63, Application US/08306063
; Patent No. 5634435
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: 5-Ethylpurylskimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; STREET: 700 Chesterfield Village Parkway

```

```

; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/306,063
; FILING DATE: 13-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,611
; FILING DATE: 28-AUG-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(10660)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6099
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 426 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-306-063-63

Query Match 0.88; Score 60; DB 1; Length 426;
Best Local Similarity 47.8%; Pred. No. 78;
Matches 11; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Q7 2028 ENPIGGEKFGPPVTKH:IPCTA 2050
Db 111 EYMLGGEPMEERPIGHLVDCLA 133

RESULT 19
US-08-833-485-63
; Sequence 63, Application US/08833485
; Patent No. 5804425
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: 5-Ethylpurylskimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,485
; FILING DATE: 07-APR-1997

```



```

1 CLASSIFICATION: 4.6
2 PRIOR APPLICATION DATA:
3 APPLICATION NUMBER: US 07/746,411
4 FILING DATE: 28-AUG-1991
5 CLASSIFICATION: 4.6
6 PRIOR APPLICATION DATA:
7 APPLICATION NUMBER: US 07/746,411
8 FILING DATE: 28-AUG-1991
9 CLASSIFICATION: 4.6
10 PRIOR APPLICATION DATA:
11 APPLICATION NUMBER: US 07/746,411
12 FILING DATE: 28-AUG-1991
13 CLASSIFICATION: 4.6
14 PRIOR APPLICATION DATA:
15 APPLICATION NUMBER: US 07/746,411
16 FILING DATE: 28-AUG-1991
17 CLASSIFICATION: 4.6
18 PRIOR APPLICATION DATA:
19 APPLICATION NUMBER: US 07/746,411
20 FILING DATE: 28-AUG-1991
21 CLASSIFICATION: 4.6
22 PRIOR APPLICATION DATA:
23 APPLICATION NUMBER: US 07/746,411
24 FILING DATE: 28-AUG-1991
25 CLASSIFICATION: 4.6
26 PRIOR APPLICATION DATA:
27 APPLICATION NUMBER: US 07/746,411
28 FILING DATE: 28-AUG-1991
29 CLASSIFICATION: 4.6
30 PRIOR APPLICATION DATA:
31 APPLICATION NUMBER: US 07/746,411
32 FILING DATE: 28-AUG-1991
33 CLASSIFICATION: 4.6
34 PRIOR APPLICATION DATA:
35 APPLICATION NUMBER: US 07/746,411
36 FILING DATE: 28-AUG-1991
37 CLASSIFICATION: 4.6
38 PRIOR APPLICATION DATA:
39 APPLICATION NUMBER: US 07/746,411
40 FILING DATE: 28-AUG-1991
41 CLASSIFICATION: 4.6
42 PRIOR APPLICATION DATA:
43 APPLICATION NUMBER: US 07/746,411
44 FILING DATE: 28-AUG-1991
45 CLASSIFICATION: 4.6
46 PRIOR APPLICATION DATA:
47 APPLICATION NUMBER: US 07/746,411
48 FILING DATE: 28-AUG-1991
49 CLASSIFICATION: 4.6
50 PRIOR APPLICATION DATA:
51 APPLICATION NUMBER: US 07/746,411
52 FILING DATE: 28-AUG-1991
53 CLASSIFICATION: 4.6
54 PRIOR APPLICATION DATA:
55 APPLICATION NUMBER: US 07/746,411
56 FILING DATE: 28-AUG-1991
57 CLASSIFICATION: 4.6
58 PRIOR APPLICATION DATA:
59 APPLICATION NUMBER: US 07/746,411
60 FILING DATE: 28-AUG-1991
61 CLASSIFICATION: 4.6
62 PRIOR APPLICATION DATA:
63 APPLICATION NUMBER: US 07/746,411
64 FILING DATE: 28-AUG-1991
65 CLASSIFICATION: 4.6
66 PRIOR APPLICATION DATA:
67 APPLICATION NUMBER: US 07/746,411
68 FILING DATE: 28-AUG-1991
69 CLASSIFICATION: 4.6
70 PRIOR APPLICATION DATA:
71 APPLICATION NUMBER: US 07/746,411
72 FILING DATE: 28-AUG-1991
73 CLASSIFICATION: 4.6
74 PRIOR APPLICATION DATA:
75 APPLICATION NUMBER: US 07/746,411
76 FILING DATE: 28-AUG-1991
77 CLASSIFICATION: 4.6
78 PRIOR APPLICATION DATA:
79 APPLICATION NUMBER: US 07/746,411
80 FILING DATE: 28-AUG-1991
81 CLASSIFICATION: 4.6
82 PRIOR APPLICATION DATA:
83 APPLICATION NUMBER: US 07/746,411
84 FILING DATE: 28-AUG-1991
85 CLASSIFICATION: 4.6
86 PRIOR APPLICATION DATA:
87 APPLICATION NUMBER: US 07/746,411
88 FILING DATE: 28-AUG-1991
89 CLASSIFICATION: 4.6
90 PRIOR APPLICATION DATA:
91 APPLICATION NUMBER: US 07/746,411
92 FILING DATE: 28-AUG-1991
93 CLASSIFICATION: 4.6
94 PRIOR APPLICATION DATA:
95 APPLICATION NUMBER: US 07/746,411
96 FILING DATE: 28-AUG-1991
97 CLASSIFICATION: 4.6
98 PRIOR APPLICATION DATA:
99 APPLICATION NUMBER: US 07/746,411
100 FILING DATE: 28-AUG-1991

```

```

1 Query Match: 0.68; Score 63; Db 4; Length 426;
2 Best Local Similarity: 47.88; Prod. No. 78;
3 Matches: 11; Conservative: 4; Mismatches: 8; Indels: 0; Gaps: 0;
4
5 QY 2028 ENPLOSPEYGFVTRPIGCLIA 2050
6 Db 111 EYMLOGPRMEERPIGHLVLA 133
7
8 RESULT 21
9 5224391-9
10 Patent No. 5224391
11 APPLICANT: SHEN, DONALD M.; DOUGARD, PAUL F.
12 TITLE OF INVENTION: INHIBITORS OF HERPES SIMPLEX VIRUS
13 REPLICATION
14 NUMBER OF SEQUENCES: 9
15 CURRENT APPLICATION DATA:
16 APPLICATION NUMBER: US/97/482,634
17 FILING DATE: 21-FEB-1990
18 SEQ ID NO: 1
19 LENGTH: 488
20 5224391-9
21
22 Query Match: 0.68; Score 60; Db 6; Length 488;
23 Best Local Similarity: 57.98; Prod. No. 99;
24 Matches: 11; Conservative: 5; Mismatches: 5; Indels: 0; Gaps: 0;
25
26 QY 101 LSPYFLKPAQKLEWLIH 119
27 Db 409 VSAVFLKPKKICLIWLIH 327
28
29 RESULT 22
30 US-08-570-311-2
31 Sequence 2, Application US/08570411
32 Patent No. 5824791
33 GENERAL INFORMATION:
34 APPLICANT: Propulse-Pax, Ann
35 APPLICANT: Tumwasorn, Somying
36 APPLICANT: Lepine, Guyline
37 APPLICANT: Han, Naiming
38 APPLICANT: Lantiz, Marilyn
39 APPLICANT: Patti, Joseph
40 TITLE OF INVENTION: Cloned Porphyromonas gingivalis genes
41 TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
42 NUMBER OF SEQUENCES: 29
43 CORRESPONDENCE ADDRESS:
44 ADDRESSEE: Ted W. Whitlock
45 STREET: 2421 N.W. 41st Street, Suite A 1
46 CITY: Gainesville
47 STATE: FL
48 COUNTRY: USA

```

21P: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311

FILING DATE: 08-SEP-1988
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640

FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UFI5.C3
TELEPHONE: (904) 375,9100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 497 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-2

Query Match 0.6%; Score 60; DB 2; Length 497;
Best Local Similarity 49.0%; Pred. No. 99;
Matches 14; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 1668 QTAATKILCKNGAENPFFVVLXTAVKLIAP 1702
DB 335 EAAFTKLEETIGSDKAPMNLVRSQVKLPAP 369

RESULT 24
US-08-453-485-2
Sequence 2, Application US/08/353,485
Patent No. 5830710
GENERAL INFORMATION:
APPLICANT: Prodisks-Prox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guyolaine
APPLICANT: Han, Naimeyn
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640

FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UFI5.C2
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 497 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-353-485-2

Query Match 0.6%; Score 60; DB 2; Length 497;
Best Local Similarity 49.0%; Pred. No. 99;
Matches 14; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 1668 QTAATKILCKNGAENPFFVVLXTAVKLIAP 1702
DB 335 EAAFTKLEETIGSDKAPMNLVRSQVKLPAP 369

RESULT 24
US-08-460-673-2
Sequence 2, Application US/08/360,673
Patent No. 5679544
GENERAL INFORMATION:
APPLICANT: Fleet, Reinhard
APPLICANT: Fournier, Alain
APPLICANT: Yeh, Patricia
TITLE OF INVENTION: MODIFIED KLUYVEROMYCES YEASTS, THEIR
PREPARATION AND USE
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Rd 3043
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19002
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/360,673
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 97/11992, 00623
FILING DATE: 23-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 92/07785
FILING DATE: 25-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Julie K.
REGISTRATION NUMBER: 38,619
REFERENCE/DOCKET NUMBER: ST92040-US
TELECOMMUNICATION INFORMATION:

TELEPHONE: (61-2)54-0000
 TELEFAX: (61-2)44-6000
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE: HSAV171313
 LENGTH: 661 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORGANISM: HSA

Query Match: 0.6% Score 60; DB 1; Length 661

Best Local Similarity: 40.0% Ident. No. 26; 26

MATCHES: 14; Conservation: 6; Mismatches: 21; Indels: 0; Gaps: 0

QY 1145 KKNKHAQVNSVKEKSNVAI VRIEL 1145

1 1111111 1111111

IR 278 NQDQDQAGLGGVAVAFNAVAKAV 278

QY 1145

US-08-570-411-14

Sequence 14: Application US/070670311

Patent No. 5824791

GENERAL INFORMATION:

APPLICANT: Promoliske-Pox, Ann

APPLICANT: Tommasoni, Somylin

APPLICANT: Lepino, Gyllatino

APPLICANT: Han, Naimio

APPLICANT: Lantz, Marlyu

APPLICANT: Patti, Joseph

TITLE OF INVENTION: Cloned Paraphromonas oimodialis genes

TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ted W. Whitlock

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/070670, 411

FILING DATE:

CLASSIFICATION: 424

APPLICATION DATA:

APPLICATION NUMBER: US 07/241,640

FILING DATE: 08-SEP-1988

ATTORNEY/AGENT INFORMATION:

NAME: Whitlock, Ted W.

REGISTRATION NUMBER: 46,965

REFERENCE/DOCKET NUMBER: 0015,03

TELECOMMUNICATION INFORMATION:

TELEPHONE: (954) 375-8100

TELEFAX: (904) 472-5800

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 779 amino acids

TYPE: AMINO ACID

STANDARDNESS: UNKNOWN

TOPOLOGY: unknown

MOLECULE TYPE: protein

US-08-570-411-14

Query Match: 0.6% Score 60; DB 1; Length 779

Best Local Similarity: 40.0% Ident. No. 26; 26

MATCHES: 14; Conservation: 6; Mismatches: 21; Indels: 0; Gaps: 0

QY 1145 KKNKHAQVNSVKEKSNVAI VRIEL 1145

1 1111111 1111111

IR 278 NQDQDQAGLGGVAVAFNAVAKAV 278

QY 1145

US-08-570-411-14

Sequence 14: Application US/070670311

Patent No. 5824791

GENERAL INFORMATION:

APPLICANT: Promoliske-Pox, Ann

APPLICANT: Tommasoni, Somylin

RESULT 26

US-08-570-411-14

Sequence 14: Application US/0670311

Patent No. 5824791

GENERAL INFORMATION:

APPLICANT: Promoliske-Pox, Ann

APPLICANT: Tommasoni, Somylin

APPLICANT: Lepino, Gyllatino

APPLICANT: Han, Naimio

APPLICANT: Lantz, Marlyu

APPLICANT: Patti, Joseph

TITLE OF INVENTION: Cloned Paraphromonas oimodialis genes

TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ted W. Whitlock

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/070670, 411

FILING DATE:

CLASSIFICATION: 424

APPLICATION DATA:

APPLICATION NUMBER: US 07/241,640

FILING DATE: 08-SEP-1988

ATTORNEY/AGENT INFORMATION:

NAME: Whitlock, Ted W.

REGISTRATION NUMBER: 46,965

REFERENCE/DOCKET NUMBER: 0015,03

TELECOMMUNICATION INFORMATION:

TELEPHONE: (954) 375-8100

TELEFAX: (904) 472-5800

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 2628 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-570-411-14

Query Match:

Best Local Similarity: 40.0% Ident. No. 26; 26

MATCHES: 14; Conservation: 6; Mismatches: 21; Indels: 0; Gaps: 0

QY 1698 QYALVILKLEKKNKHAQVNSVKEKSNVAI VRIEL 1702

1 1111111 1111111

IR 335 EAANFLIKLLEKKNKHAQVNSVKEKSNVAI VRIEL 369

RESULT 27

US-08-444 818-10

Sequence 14: Application US/06044-818

Patent No. 6150067

GENERAL INFORMATION:

APPLICANT: Chien, David Y.

```

? APPLICANT: Rutter, William J.
? TITLE OF INVENTION: NANV Diagnostics and Vaccines
? NUMBER OF SEQUENCES: 777
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Chiron Corporation
? STREET: 4560 Horton Street
? CITY: Emeryville
? STATE: CA
? COUNTRY: USA
? ZIP: 94608-2916
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: IBM PC compatible
? SOFTWARE: Patent In Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/444,818
? FILING DATE:
? CLASSIFICATION: 424
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US/08/403,590
? FILING DATE: 14-MAR-1995
? ATTORNEY/AGENT INFORMATION:
? NAME: Harbin, Alisa A.
? REGISTRATION NUMBER: 33,895
? REFERENCE/DOCKET NUMBER: 0110,002
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (508)359-3876
? TELEFAX: (508)359-3885
? INFORMATION FOR SEQ ID NO: 10:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 117 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-444-818-10

```

```

Query Match 0.5%; Score 59; DB 4; Length 117;
Best Local Similarity 34.5%; Pred. No. 15;
Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

```

```

QY 1821 VLIIPAIIKTYKQIEKNKNNMGPFMSIIQ 1849
I: I I I I I I I I I I I I I I I I
DB 57 VIAPAVOTNWKLETFWAKIMWNFISGIQ 85

```

```

RESULT 28
PCT-US92-07811-12
? Sequence 12, Application PCT/US9207813
? GENERAL INFORMATION:
? APPLICANT: LESNIEWSKI, RICHARD R.
? APPLICANT: LEUNG, TAT K.
? TITLE OF INVENTION: HEPATITIS C ASSAY
? NUMBER OF SEQUENCES: 27
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: ABBOTT LABORATORIES
? STREET: ONE ABBOTT PARK ROAD
? CITY: ABBOTT PARK
? STATE: ILLINOIS
? COUNTRY: U.S.A.
? ZIP: 60054-3500
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: IBM PC compatible
? SOFTWARE: Patent In Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: PCT/US92/07813
? FILING DATE: 19920916
? CLASSIFICATION:
? ATTORNEY/AGENT INFORMATION:
? NAME: FOREMSKI, PRISCILLA E.

```

```

? REGISTRATION NUMBER: 33,207
? REFERENCE/DOCKET NUMBER: 4767.P3.03
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 708-937-6365
? TELEFAX: 708-937-9556
? INFORMATION FOR SEQ ID NO: 12:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 117 amino acids
? TYPE: AMINO ACID
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? PCT-US92-07813-12

```

```

Query Match 0.5%; Score 59; DB 5; Length 117;
Best Local Similarity 34.5%; Pred. No. 15;
Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

```

```

QY 1821 VLIIPAIIKTYKQIEKNKNNMGPFMSIIQ 1849
I: I I I I I I I I I I I I I I I I
DB 57 VIAPAVOTNWKLETFWAKIMWNFISGIQ 85

```

```

RESULT 29
US-08-444-818-8
? Sequence 8, Application US/08444818
? Patent No. 6150087
? GENERAL INFORMATION:
? APPLICANT: Chien, David Y.
? TITLE OF INVENTION: NANV Diagnostics and Vaccines
? NUMBER OF SEQUENCES: 777
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Chiron Corporation
? STREET: 4560 Horton Street
? CITY: Emeryville
? STATE: CA
? COUNTRY: USA
? ZIP: 94608-2916
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: IBM PC compatible
? SOFTWARE: Patent In Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/444,818
? FILING DATE:
? CLASSIFICATION: 424
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US/08/403,590
? FILING DATE: 14-MAR-1995
? ATTORNEY/AGENT INFORMATION:
? NAME: Harbin, Alisa A.
? REGISTRATION NUMBER: 33,895
? REFERENCE/DOCKET NUMBER: 0110,002
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (508)359-3876
? TELEFAX: (508)359-3885
? INFORMATION FOR SEQ ID NO: 8:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 128 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-444-818-8

```

```

Query Match 0.5%; Score 59; DB 4; Length 128;
Best Local Similarity 34.5%; Pred. No. 17;
Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

```

```

QY 1821 VLIIPAIIKTYKQIEKNKNNMGPFMSIIQ 1849

```



```

1 FILING DATE: 21-AUG-1991
2 PRIOR APPLICATION DATA:
3 APPLICATION NUMBER: US 07/748,566
4 FILING DATE: 21-AUG-1991
5 PRIOR APPLICATION DATA:
6 APPLICATION NUMBER: US 07/748,566
7 FILING DATE: 21-AUG-1991
8 ATTORNEY/AGENT INFORMATION:
9 NAME: PEREMBSKI, PRISCILLA E
10 REGISTRATION NUMBER: 33,209
11 REFERENCE/AGENT NUMBER: 4044 US.P6
12 TELECOMMUNICATION INFORMATION
13 TELEPHONE: 708-947-6365
14 TELEFAX: 708-947-9556
15 INFORMATION FOR SEQ ID NO: 47:
16 SEQUENCE CHARACTERISTICS:
17 LENGTH: 594 amino acids
18 TYPE: amino acid
19 STRANDEDNESS: single
20 TOPOLOGY: linear
21 MOLECULE TYPE: peptide
22 US-09-603-611-47
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

```

1 FILING DATE: 24-AUG-1990
2 PRIOR APPLICATION DATA:
3 APPLICATION NUMBER: US 07/748,566
4 FILING DATE: 07-NOV-1990
5 PRIOR APPLICATION DATA:
6 APPLICATION NUMBER: US 07/748,561
7 FILING DATE: 21-AUG-1991
8 PRIOR APPLICATION DATA:
9 APPLICATION NUMBER: US 07/748,565
10 FILING DATE: 21-AUG-1991
11 PRIOR APPLICATION DATA:
12 APPLICATION NUMBER: US 07/748,566
13 FILING DATE: 21-AUG-1991
14 ATTORNEY/AGENT INFORMATION:
15 NAME: PEREMBSKI, PRISCILLA E
16 REGISTRATION NUMBER: 33,209
17 REFERENCE/AGENT NUMBER: 4044 US.P6
18 TELECOMMUNICATION INFORMATION
19 TELEPHONE: 708-947-6365
20 TELEFAX: 708-947-9556
21 INFORMATION FOR SEQ ID NO: 48:
22 SEQUENCE CHARACTERISTICS:
23 LENGTH: 594 amino acids
24 TYPE: amino acid
25 STRANDEDNESS: single
26 TOPOLOGY: linear
27 MOLECULE TYPE: peptide
28 US-08-867-611-48
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

APPLICATION NUMBER: US/08/646,757
 FILING DATE: 02/08/1999
 APPLICATION NUMBER: US/08/179,896
 FILING DATE: 02/08/1999
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/572,922
 FILING DATE: 24-AUG-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/614,069
 FILING DATE: 07 NOV 1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/748,561
 FILING DATE: 21-AUG-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/748,565
 FILING DATE: 21-AUG-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/748,566
 FILING DATE: 21-AUG-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: POREMSKI, PRISCILLA E
 REGISTRATION NUMBER: 33,207
 REFERENCE/DOCKET NUMBER: 4834, US, DE
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 708-937-6365
 TELEFAX: 708-937-9556
 INFORMATION FOR SEQ ID NO: 16:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 597 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-867-611-18

Query Match 0.5%; Score 59; DB 4; Length 597;
 Best Local Similarity 34.5%; Pred. No. 1.8e+02;
 Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1821 VLLPAIKKLYKQIFKMKKNDKPFMSHLLQ 1819
 I: I::: I::: I: I: I: I:

Db 401 VIAPAVQTNWQKLETFWAKHWNFISGIQ 429

RESULT 37
 PCT-US92-06965A-21
 Sequence 21, Application PC/US9206965A
 GENERAL INFORMATION:
 APPLICANT: DEVARE, SUSHIL G.
 APPLICANT: DESAI, SURESH M.
 APPLICANT: CASEY, JAMES M.
 APPLICANT: DAILEY, STEPHEN H.
 APPLICANT: DAWSON, GEORGE J.
 APPLICANT: GUTIERREZ, ROBIN A.
 APPLICANT: LESNIEWSKI, RICHARD P.
 APPLICANT: STEWART, JAMES L.
 APPLICANT: RUPPRECHT, KEVIN R.
 TITLE OF INVENTION: RECY SYNTHETIC PEPTIDE FROM HSI REGION
 NUMBER OF SEQUENCES: 35
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: ABBOTT LABORATORIES
 STREET: ONE ABBOTT PARK ROAD
 CITY: ABBOTT PARK
 STATE: ILLINOIS
 COUNTRY: U.S.
 ZIP: 60065-3500
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US92-06965A
 FILING DATE: 19920821
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: POREMSKI, PRISCILLA E.
 REGISTRATION NUMBER: 33,207
 REFERENCE/DOCKET NUMBER: 4834PC.02

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 708-937-6365
 TELEFAX: 708-937-9556
 INFORMATION FOR SEQ ID NO: 21:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 597 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US92-06965A-21

Query Match 0.5%; Score 59; DB 5; Length 597;
 Best Local Similarity 34.5%; Pred. No. 1.8e+02;
 Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1821 VLLPAIKKLYKQIFKMKKNDKPFMSHLLQ 1819
 I: I::: I::: I: I: I: I:

Db 401 VIAPAVQTNWQKLETFWAKHWNFISGIQ 429

RESULT 38
 US-08-867-611-18
 Sequence 18, Application US/08867611
 Patent No. 6172189
 GENERAL INFORMATION:
 APPLICANT: DEVARE, SUSHIL G.
 APPLICANT: DESAI, SURESH M.
 APPLICANT: CASEY, JAMES M.
 APPLICANT: DAILEY, STEPHEN H.
 APPLICANT: DAWSON, GEORGE J.
 APPLICANT: GUTIERREZ, ROBIN A.
 APPLICANT: LESNIEWSKI, RICHARD P.
 APPLICANT: STEWART, JAMES L.
 APPLICANT: RUPPRECHT, KEVIN R.
 TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
 NUMBER OF SEQUENCES: 59
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: ABBOTT LABORATORIES
 STREET: ONE ABBOTT PARK ROAD, CHAN477/AP602
 CITY: ABBOTT PARK
 STATE: IL
 COUNTRY: USA
 ZIP: 60064-3500
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/867,611
 FILING DATE: 02-JUN-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/646,757
 FILING DATE: 02-JUN-1997
 APPLICATION NUMBER: US/08/179,896
 FILING DATE: 24-AUG-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/572,822
 FILING DATE: 24-AUG-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/614,069
 FILING DATE: 07-NOV-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/748,561
 FILING DATE: 21-AUG-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/748,565
 FILING DATE: 21-AUG-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/748,566

27 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 28 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 29 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 30 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 31 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 32 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 33 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 34 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 35 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 36 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 37 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 38 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 39 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 40 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 41 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 42 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 43 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 44 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 45 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 46 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 47 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 48 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 49 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 50 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 51 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 52 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 53 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 54 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 55 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 56 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 57 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 58 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 59 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 60 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 61 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 62 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 63 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 64 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 65 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 66 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 67 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 68 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 69 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 70 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 71 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 72 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 73 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 74 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 75 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 76 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 77 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 78 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 79 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 80 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 81 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 82 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 83 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 84 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 85 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 86 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 87 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 88 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 89 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 90 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 91 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 92 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 93 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 94 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 95 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 96 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 97 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 98 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 99 1921 VILPAKREY, JEROME M. PERS. 1949
 Best Local Similarity: 41.9% (Seq. No. 1, 60-02)
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;
 100 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441

DB: 403 VIAPAVQTNWQKLETWAKIMWNI1831Q 441
 RESULT: 40
 US-08-867-611-49
 Sequence: 49, Application: US/08067611
 Patent No.: 6172189
 GENERAL INFORMATION:
 APPLICANT: DEVARE, SUSHIL G
 APPLICANT: DESAI, SURESH M
 APPLICANT: CASEY, JAMES M
 APPLICANT: DAILEY, STEPHEN F
 APPLICANT: LAWSON, GEORGE J
 APPLICANT: GUTIERREZ, ROBIN A
 APPLICANT: LESNIEWSKI, RICHARD R
 APPLICANT: STEWART, JAMES L
 APPLICANT: ROPPRECHT, KEVIN R
 TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING KITS/REAGENTS
 TITLE OF INVENTION: ANTIGENS
 NUMBER OF SEQUENCES: 59
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: ABBOTT LABORATORY FILES
 STREET: ONE ABBOTT PARK ROAD, CHAS. 07748/AD-12
 CITY: ABBOTT PARK
 STATE: IL
 COUNTRY: USA
 ZIP: 60064-3500
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/646,757
 FILING DATE: 02-JUN-1997
 CLASSIFICATION: 445
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/646,757
 FILING DATE:
 APPLICATION NUMBER: US/08/179,896
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/572,922
 FILING DATE: 24-AUG-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/614,069
 FILING DATE: 07-NOV-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/748,561
 FILING DATE: 21-AUG-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/748,565
 FILING DATE: 21-AUG-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/748,566
 FILING DATE: 21-AUG-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: POREMSKI, PRISCILLA E
 REGISTRATION NUMBER: 33,267
 REFERENCE/DOCKET NUMBER: 444,05,19
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 708-937-5365
 TELEFAX: 708-937-9556
 INFORMATION FOR SEQ ID NO: 49:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 613 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-867-611-49

us-09-603-665-5.ra1

Fri Nov 9 15:42:38 2001

Query Match 0.5%; Score 59; DB 4; Length 613;
Best Local Similarity 34.5%; Pred. No. 1.9e+02;
Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1821 VLLPAIKRYKQIPFNWPNHMCPEMSILQ 1849
L: I: : : : : I: I: I: I: I:
DB 422 VIAPAVQTNWQKLETFWAKHWNFISGIQ 450

RESULT 41
US-08-444-818-148
Sequence 148; Application US/08444818
Patent No. 6150087
GENERAL INFORMATION:
APPLICANT: Chien, David Y.
APPLICANT: Butler, William J.
TITLE OF INVENTION: NANIV Diagnostics and Vaccines
NUMBER OF SEQUENCES: 777
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,818
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/403,540
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Harbin, Alisa A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 9110.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508)359-3876
TELEFAX: (508)359-3885
INFORMATION FOR SEQ ID NO: 148:
SEQUENCE CHARACTERISTICS:
LENGTH: 739 amino acids
TYPE: amino acid
TOPOLOGY: Linear
MOLECULE TYPE: protein
US-08-444-818-148

Query Match 0.5%; Score 59; DB 4; Length 739;
Best Local Similarity 34.5%; Pred. No. 2.5e+02;
Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1821 VLLPAIKRYKQIPFNWPNHMCPEMSILQ 1849
L: I: : : : : I: I: I: I: I:
DB 554 VIAPAVQTNWQKLETFWAKHWNFISGIQ 582

RESULT 42
US-08-867-611-4
Sequence 4; Application US/08867611
Patent No. 6172189
GENERAL INFORMATION:
APPLICANT: DEVARE, SUSHIL G
APPLICANT: DESAI, SURESH M
APPLICANT: CASEY, JAMES M
APPLICANT: BALLEW, STEPHEN H
APPLICANT: DAWSON, GEORGE J
APPLICANT: GUTIERREZ, ROBIN A

APPLICANT: LESNIEWSKI, RICHARD R
APPLICANT: STEWART, JAMES L
APPLICANT: RUPPRECHT, KEVIN R
TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES
STREET: ONE ABBOTT PARK ROAD, CHADWY/AP6D2
CITY: ABBOTT PARK
STATE: IL USA
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/867,611
FILING DATE: 02-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/646,757
FILING DATE:
APPLICATION NUMBER: US/08/179,896
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/572,922
FILING DATE: 24-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/614,069
FILING DATE: 07-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/748,561
FILING DATE: 21-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/748,565
FILING DATE: 21-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/748,566
FILING DATE: 21-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: FOREMSKI, PRISCILLA E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 4834 US.P6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-937-9556
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 781 amino acids
TYPE: amino acid
TOPOLOGY: Linear
MOLECULE TYPE: protein
US-08-867-611-4

Query Match 0.5%; Score 59; DB 4; Length 781;
Best Local Similarity 34.5%; Pred. No. 2.7e+02;
Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1821 VLLPAIKRYKQIPFNWPNHMCPEMSILQ 1849
L: I: : : : : I: I: I: I: I:
DB 585 VIAPAVQTNWQKLETFWAKHWNFISGIQ 613

RESULT 43
PCT-US92-06965A-9
Sequence 9; Application PC/TUS9206965A
GENERAL INFORMATION:
APPLICANT: DEVARE, S.
APPLICANT: DESAI, S.

1 FILING DATE: 21-AUG-1991
2 PRIOR APPLICATION DATA:
3 APPLICATION NUMBER: US 07/748,566
4 FILING DATE: 21-AUG-1991
5 ATTORNEY/AGENT INFORMATION:
6 NAME: POREMBSKI, PRISCILLA E
7 REGISTRATION NUMBER: 33,207
8 REFERENCE/DOCKET NUMBER: 4844.US.P6
9 TELECOMMUNICATION INFORMATION:
10 TELEPHONE: 708-937-6365
11 TELEFAX: 708-937-9556
12 INFORMATION FOR SEQ ID NO: 52:
13 SEQUENCE CHARACTERISTICS:
14 LENGTH: 971 amino acids
15 TYPE: amino acid
16 STRANDEDNESS: single
17 TOPOLOGY: linear
18 MOLECULE TYPE: peptide
19 US-08-867-611-52

Query Match 0.5%; Score 59; DB 4; Length 971;
Best Local Similarity 34.5%; Pred. No. 3.8e+02;
Matches 10; Conservative 9; Mismatches 16; Indels 3; Gaps 0;

QY 1821 VLLPAIKKTKYQJEXNKNHMGPFMSIIQ 1849
1 : | | : : : | | | : | : |
Db 780 VIAFAVQTNQKLETFWAKIMWNFISGIQ 808

RESULT 46
US-08-867-611-53
1 Sequence 53, Application US/08867611
2 Patent No. 6172189
3 GENERAL INFORMATION:
4 APPLICANT: DEVARE, SUSHIL G
5 APPLICANT: DESAI, SURESH M
6 APPLICANT: CASEY, JAMES M
7 APPLICANT: DAILEY, STEPHEN H
8 APPLICANT: DAWSON, GEORGE J
9 APPLICANT: GUTIERREZ, ROBIN A
10 APPLICANT: LESNIEWSKI, RICHARD R
11 APPLICANT: STEWART, JAMES L
12 APPLICANT: RUPPRECHT, KEVIN R
13 TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
14 NUMBER OF SEQUENCES: 59
15 CORRESPONDENCE ADDRESS:
16 ADDRESSEE: ABBOTT LABORATORIES
17 STREET: ONE ABBOTT PARK ROAD, CHADDY77/AP602
18 CITY: ABBOTT PARK
19 STATE: IL
20 COUNTRY: USA
21 ZIP: 60064-3500
22 COMPUTER READABLE FORM:
23 MEDIUM TYPE: Floppy disk
24 OPERATING SYSTEM: PC-DOS/MS-DOS
25 SOFTWARE: Patent In Release #1.0, Version #1.25
26 CURRENT APPLICATION DATA:
27 FILING DATE: 02-JUN-1997
28 CLASSIFICATION: 435
29 PRIOR APPLICATION DATA:
30 APPLICATION NUMBER: US/08/646,757
31 FILING DATE:
32 FILING DATE:
33 PRIOR APPLICATION DATA:
34 APPLICATION NUMBER: US/07/573,822
35 FILING DATE: 24-AUG-1990
36 PRIOR APPLICATION DATA:
37 APPLICATION NUMBER: US 07/614,069

1 FILING DATE: 07-NOV-1990
2 PRIOR APPLICATION DATA:
3 APPLICATION NUMBER: US 07/748,561
4 FILING DATE: 21-AUG-1991
5 PRIOR APPLICATION DATA:
6 APPLICATION NUMBER: US 07/748,565
7 FILING DATE: 21-AUG-1991
8 PRIOR APPLICATION DATA:
9 APPLICATION NUMBER: US 07/748,566
10 FILING DATE: 21-AUG-1991
11 ATTORNEY/AGENT INFORMATION:
12 NAME: POREMBSKI, PRISCILLA E
13 REGISTRATION NUMBER: 33,207
14 REFERENCE/DOCKET NUMBER: 4844.US.P6
15 TELECOMMUNICATION INFORMATION:
16 TELEPHONE: 708-937-6365
17 TELEFAX: 708-937-9556
18 INFORMATION FOR SEQ ID NO: 53:
19 SEQUENCE CHARACTERISTICS:
20 LENGTH: 973 amino acids
21 TYPE: amino acid
22 STRANDEDNESS: single
23 TOPOLOGY: linear
24 MOLECULE TYPE: peptide
25 US-08-867-611-53

Query Match 0.5%; Score 59; DB 4; Length 973;
Best Local Similarity 34.5%; Pred. No. 3.8e+02;
Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1821 VLLPAIKKTKYQJEXNKNHMGPFMSIIQ 1849
1 : | | : : : | | | : | : |
Db 782 VIAFAVQTNQKLETFWAKIMWNFISGIQ 810

RESULT 47
US-08-867-611-54
1 Sequence 54, Application US/08867611
2 Patent No. 6172189
3 GENERAL INFORMATION:
4 APPLICANT: DEVARE, SUSHIL G
5 APPLICANT: DESAI, SURESH M
6 APPLICANT: CASEY, JAMES M
7 APPLICANT: DAILEY, STEPHEN H
8 APPLICANT: DAWSON, GEORGE J
9 APPLICANT: GUTIERREZ, ROBIN A
10 APPLICANT: LESNIEWSKI, RICHARD R
11 APPLICANT: STEWART, JAMES L
12 APPLICANT: RUPPRECHT, KEVIN R
13 TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
14 NUMBER OF SEQUENCES: 59
15 CORRESPONDENCE ADDRESS:
16 ADDRESSEE: ABBOTT LABORATORIES
17 STREET: ONE ABBOTT PARK ROAD, CHADDY77/AP602
18 CITY: ABBOTT PARK
19 STATE: IL
20 COUNTRY: USA
21 ZIP: 60064-3500
22 COMPUTER READABLE FORM:
23 MEDIUM TYPE: Floppy disk
24 OPERATING SYSTEM: IBM PC compatible
25 SOFTWARE: Patent In Release #1.0, Version #1.25
26 CURRENT APPLICATION DATA:
27 APPLICATION NUMBER: US/08/667,611
28 FILING DATE: 02-JUN-1997
29 CLASSIFICATION: 435
30 PRIOR APPLICATION DATA:
31 APPLICATION NUMBER: US/08/646,757
32 FILING DATE:
33 FILING DATE:
34 APPLICATION NUMBER: US/08/179,896

```

; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,624
; FILING DATE: 24 AUG 1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/714,069
; FILING DATE: 07 NOV 1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,621
; FILING DATE: 24 AUG 1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,605
; FILING DATE: 21 AUG 1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,604
; FILING DATE: 21 AUG 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: PEREMSKI, PRISCILLA E
; REGISTRATION NUMBER: 34,257
; REFERENCE/LOCKET NUMBER: 4454, US 15
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510 601-2702
; TELEFAX: 510 655-4542
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 942 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; STRAIN/GENE: Stable
; MOLECULE TYPE: protein
; MISMATCHES: 10; Indels: 0; Gaps: 0
; 5' 603 611 54
; 5' 603 611 54

Query Match: 0.58, Score 59, DB 1, Length 1021
Best Local Similarity: 34.58, Pred. No. 4, loc-02
Matches: 10, Conservative: 9, Mismatches: 10; Indels: 0; Gaps: 0

; 1221 VLLPAVETFEYTEREKWYHMHPTASTIQ 1849
; 1 1111 1111 1111
; 713 VIAVAVQINWKLKLEFWAKHMMNFISGIG 741

RESULT 49
US-08-440-519-12
; Sequence 12, Application US/0844 519
; Patent No. 5712087
; GENERAL INFORMATION:
; APPLICANT: Boudhott, Michael
; APPLICANT: Choo, Qui-Lim
; APPLICANT: Kuo, George
; TITLE OF INVENTION: Combinations of Hepatitis B virus (HBV)
; TITLE OF INVENTION: Anticaps for Use in Immunoassays for Anti HBV Antibodies
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: P.O. Box 8097 (Int. Prop. R-440)
; CITY: Emeryville
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC DOS/MS DOS
; SOFTWARE: Patent In Release #1.0, Version #1.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08440,519
; FILING DATE: 12 MAY 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/910,790
; FILING DATE: 07 JUL 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Blackburn Esq., Robert P.
; REGISTRATION NUMBER: 40,447
; REFERENCE/LOCKET NUMBER: 0 01,002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2702
; TELEFAX: (510) 655-4542
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1021 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-440-519-12

Query Match: 0.58, Score 59, DB 1, Length 1021
Best Local Similarity: 34.58, Pred. No. 4, loc-02
Matches: 10, Conservative: 9, Mismatches: 10; Indels: 0; Gaps: 0

; 1221 VLLPAVETFEYTEREKWYHMHPTASTIQ 1849
; 1 1111 1111 1111

```

```

b  713 VIAPAVQTRNWKLELTHWAKRRTFSSQ 741

RESULT 50
Sequence 2, Application US/09648738
Patent No. 6207374
GENERAL INFORMATION:
  APPLICANT: Sampson et al.
  TITLE OF INVENTION: Tuberculous Sclerosis 2 Gene and Uses
  TITLE OF INVENTION: Thercoet
  NUMBER OF SEQUENCES: 4
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Kathleen M. Williams, Banner & Witcoff,
  ADDRESSEE: Ltd.
  STREET: One Financial Center
  CITY: Boston
  STATE: Massachusetts
  COUNTRY: USA
  ZIP: 02111
COMPUTER READABLE FORM:
  MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: Wordperfect 6.1a
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/040.738
  FILING DATE: Concurrently herewith
  CLASSIFICATION:
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: GB9326470.3
  FILING DATE: 24-December-1993
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: GB9411900.5
  FILING DATE: 14-June-1994
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: PC1/GB94/02823
  FILING DATE: 23-December-1994
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: GB/952.426
  FILING DATE: 30-May-1996
  ATTORNEY/AGENT INFORMATION:
  NAME: Williams, Kathleen M.
  REGISTRATION NUMBER: 34,380
  REFERENCE/DOCUMENT NUMBER: 326573963
  TELECOMMUNICATION INFORMATION:
  TELEPHONE: 617-345-9100
  TELEFAX: 617-345-9111
  INFORMATION FOR SEQ ID NO: 2:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 1784 amino acids
  TYPE: amino acid
  TOPOLOGY: linear
  MOLECULE TYPE: protein
US-09_040-738-2

Query Match 0.5%; Score 59; bh 4; Length 1784;
Best Local Similarity 25.8%; Pred.No. 9.7e+02;
Matches 17; Conservative 17; Mismatches 32; Indels 0; Gaps 0;

QY 1324 MSSTFPRANVHLEDTSEGVIRIVKRVIAI LQSTGSLSTFVSFNVFTVVFISV 1382
DB 172 LSSERFLVILNLIKFKNSYILIRVIAHVVLMICLLCVRIASSVDIEVSLQVLDVAVVCYNCL 231
QY 1383 FVDALP 1388
DB 222 FAEELP 237

RESULT 51
US-09-652-426A-2

```



```

QY 541 FEIFKEHFSSEVTISNLNLF 561
| : : : | : : |
Db 681 FAMWKITYKDVQLRRKLEFF 701

RESULT 55
US-08-461-503-2
; Sequence 2, Application US/08461503
; Patent No. 5844302
; GENERAL INFORMATION:
; APPLICANT: Racanelli, Vincent
; APPLICANT: Tatem, Joanne M.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES
; TITLE OF INVENTION: FROM CUNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,503
; FILING DATE: 5-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-D-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-461-503-2

Query Match 0.5%; Score 59; DB 2; Length 2206;
Best Local Similarity 21.0%; Prod No. 1.3e+03;
Matches 17; Conservative 17; Mismatches 47; Indels 0; Gaps 0;

QY 481 TSLMLSHLPLAVRILAMNHLKIMKTSFGVDESFIKAVLARLGDNDIVVLSAISA 540
| : : : | : : |
Db 621 TAVETGATNPLASDTVGTTHVVGRRPSESTIIPSPAPACACVATIEVDNEQPTTRAKL 680

QY 541 FEIFKEHFSSEVTISNLNLF 561
| : : : | : : |
Db 681 FAMWKITYKDVQLRRKLEFF 701

RESULT 57
US-08-444-818-66
; Sequence 66, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:

```



```

: APPLICATION NUMBER: 09/09444818
: FILING DATE: 14 MAR 1996
: ATT-PRY/AGENT INFORMATION
: NAME: BATHIA, ALLISA A.
: PERSISTENT NUMBER: 33,896
: REFERENCE/SEQ ID NO: 1
: TELECOMMUNICATION INFORMATION
: TELEPHONE: (508) 459-8076
: TELEFAX: (508) 459-8086
: INFORMATION FOR SEQ ID NO: 66:
: SEQUENCE CHARACTERISTICS
: LENGTH: 2549 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US 09-444 818-66

Query Match
Best Local Similarity 34.5% Score 59; DB 4; Length 2549;
Matches 11; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1621 VLLEAIKRIYKQIEKKNKNGPFMSILQ 1649
: 11: 11: 11: 11: 11: 11: 11: 11: 11: 11:
DB 1129 VIAPVQINWQKLETEWAKRWNNFISG1Q 1424

RESULT 59
PCT-US95-06722 12
: Sequence 12; Application PCT/US95/06722
: GENERAL INFORMATION
: APPLICANT
: TITLE OF INVENTION: Immunosuppressant Target Proteins
: NUMBER OF SEQUENCES: 25
: COMPUTER READABLE FORM
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: ASCII (Text)
: CURRENT APPLICATION DATA
: APPLICATION NUMBER: PCT/US95/06722
: FILING DATE
: CLASSIFICATION
: PRIOR APPLICATION DATA
: APPLICATION NUMBER: US 08/250,795
: FILING DATE: 27 MAY 1994
: PRIOR APPLICATION DATA
: APPLICATION NUMBER: US 08/250,795
: FILING DATE: 20-DEC-1994
: INFORMATION FOR SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS
: LENGTH: 2549 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: PCT-US95-06722-12

Query Match
Best Local Similarity 31.4% Score 59; DB 5; Length 2549;
Matches 11; Conservative 12; Mismatches 12; Indels 0; Gaps 0;

QY 2070 LKRTPESSPEVPAALIVIALAFIEFNVILIP 2104
: 11: 11: 11: 11: 11: 11: 11: 11: 11: 11:
DB 1159 IVKTLQSPELESTAMULSSIVFGLGKKYQIFP 1194

RESULT 60
US-08-444-818-89
: Sequence 27; Application US/08444818
: Patent No. 6150087
: GENERAL INFORMATION
: APPLICANT: Chien, David Y.
: ATT-PRY/AGENT INFORMATION
: NAME: BATHIA, ALLISA A.
: PERSISTENT NUMBER: 33,896
: REFERENCE/SEQ ID NO: 1
: TELECOMMUNICATION INFORMATION
: TELEPHONE: (508) 459-8076
: TELEFAX: (508) 459-8086
: INFORMATION FOR SEQ ID NO: 77:
: SEQUENCE CHARACTERISTICS
: LENGTH: 2400 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US 08-444 818-77

Query Match
Best Local Similarity 34.5% Score 59; DB 4; Length 2400;
Matches 11; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

```

```

: APPLICATION NUMBER: 09/09444818
: FILING DATE: 14 MAR 1996
: ATT-PRY/AGENT INFORMATION
: NAME: BATHIA, ALLISA A.
: PERSISTENT NUMBER: 33,896
: REFERENCE/SEQ ID NO: 1
: TELECOMMUNICATION INFORMATION
: TELEPHONE: (508) 459-8076
: TELEFAX: (508) 459-8086
: INFORMATION FOR SEQ ID NO: 66:
: SEQUENCE CHARACTERISTICS
: LENGTH: 2549 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US 09-444 818-66

Query Match
Best Local Similarity 34.5% Score 59; DB 4; Length 2549;
Matches 11; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1621 VLLEAIKRIYKQIEKKNKNGPFMSILQ 1649
: 11: 11: 11: 11: 11: 11: 11: 11: 11: 11:
DB 1295 VIAPVQINWQKLETEWAKRWNNFISG1Q 1424

RESULT 59
PCT-US95-06722 12
: Sequence 12; Application PCT/US95/06722
: GENERAL INFORMATION
: APPLICANT
: TITLE OF INVENTION: Immunosuppressant Target Proteins
: NUMBER OF SEQUENCES: 25
: COMPUTER READABLE FORM
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: ASCII (Text)
: CURRENT APPLICATION DATA
: APPLICATION NUMBER: PCT/US95/06722
: FILING DATE
: CLASSIFICATION
: PRIOR APPLICATION DATA
: APPLICATION NUMBER: US 08/250,795
: FILING DATE: 27 MAY 1994
: PRIOR APPLICATION DATA
: APPLICATION NUMBER: US 08/250,795
: FILING DATE: 20-DEC-1994
: INFORMATION FOR SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS
: LENGTH: 2549 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: PCT-US95-06722-12

Query Match
Best Local Similarity 31.4% Score 59; DB 5; Length 2549;
Matches 11; Conservative 12; Mismatches 12; Indels 0; Gaps 0;

QY 2070 LKRTPESSPEVPAALIVIALAFIEFNVILIP 2104
: 11: 11: 11: 11: 11: 11: 11: 11: 11: 11:
DB 1159 IVKTLQSPELESTAMULSSIVFGLGKKYQIFP 1194

RESULT 60
US-08-444-818-89
: Sequence 27; Application US/08444818
: Patent No. 6150087
: GENERAL INFORMATION
: APPLICANT: Chien, David Y.
: ATT-PRY/AGENT INFORMATION
: NAME: BATHIA, ALLISA A.
: PERSISTENT NUMBER: 33,896
: REFERENCE/SEQ ID NO: 1
: TELECOMMUNICATION INFORMATION
: TELEPHONE: (508) 459-8076
: TELEFAX: (508) 459-8086
: INFORMATION FOR SEQ ID NO: 77:
: SEQUENCE CHARACTERISTICS
: LENGTH: 2400 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US 08-444 818-77

Query Match
Best Local Similarity 34.5% Score 59; DB 4; Length 2400;
Matches 11; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

```

APPLICATION NUMBER: US 603,403,540
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Harbin, Alisa A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0110.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508)359-3876
TELEFAX: (508)359-3885
INFORMATION FOR SEQ ID NO: 99
SEQUENCE CHARACTERISTICS:
LENGTH: 2772 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-444-818-89

Query Match 0.5%; Score 59; DB 4; Length 2772;
Best Local Similarity 34.5%; Pred. No. 1.9e+03;
Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1821 VILPAKRTYKLEKRWKNIHGFMSILQ 1849
1: 111 111 11 11 11 11
DB 1631 VLAFAVQTNWQKLETFWAKHMMNFISGIQ 1659
1: 111 111 11 11 11 11

RESULT 61
US-08-466-975A-23
Sequence 23, Application US/08/466-975A
Patent No. 5910404
GENERAL INFORMATION:
APPLICANT: DELEYS, ROBERT J
APPLICANT: POLLET, DIRK
APPLICANT: MAERTENS, GERT
TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
TITLE OF INVENTION: ANTIBODIES TO HEPATITIS C VIRUS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,975A
FILING DATE: 14-OCT-1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/391,671
FILING DATE: 14-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100

INFORMATION FOR SEQ ID NO: 23
SEQUENCE CHARACTERISTICS:
LENGTH: 2894 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-466-975A-23

Query Match 0.5%; Score 59; DB 2; Length 2894;
Best Local Similarity 34.5%; Pred. No. 2e+03;
Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1821 VILPAKRTYKLEKRWKNIHGFMSILQ 1849
1: 111 111 11 11 11 11
DB 1745 VIAPAVQTNWQKLETFWAKHMMNFISGIQ 1773
1: 111 111 11 11 11 11

RESULT 62
US-08-391-671A-23
Sequence 23, Application US/08/391671A
Patent No. 5922532
GENERAL INFORMATION:
APPLICANT: DELEYS, ROBERT J
APPLICANT: POLLET, DIRK
APPLICANT: MAERTENS, GERT
TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
TITLE OF INVENTION: ANTIBODIES TO HEPATITIS C VIRUS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/391,671A
FILING DATE: 21-SEP-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/920,286
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/EP91/02409
FILING DATE: 13-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 6012441.2
FILING DATE: 14-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100
INFORMATION FOR SEQ ID NO: 23
SEQUENCE CHARACTERISTICS:
LENGTH: 2894 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO

Query Match
Best Local Similarity: 44.58; Score: 59; ID: 2; Length: 2694;
Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;

QY 1821 VLLPAIKFYVGFNFNMHMFNSIQ 1849
DE 1745 VIAVAVQTNWQKLTFWAKHNRFSGIQ 1773

RESULT 64
US-09-275-265-23
Sequence 23, Application US/09275265
Patent No. 6,297,761
GENERAL INFORMATION:
APPLICANT: DELVEY, ROBERT J.
APPLICANT: POLLET, DIRK
APPLICANT: MAERTENS, GEERT
TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
TITLE OF INVENTION: ANTIBODIES TO HEPATITIS C VIRUS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: RIXER & VANERHUYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.40
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/275,265
FILING DATE: 14-DEC-1999
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/98/091,671
FILING DATE: 21-FEB-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/97/420,286
FILING DATE: 14-DEC-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: WO 97/0791/02409
FILING DATE: 13-DEC-1991
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: EP 96134241.2
FILING DATE: 14-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: SADDFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/BOOKLET NUMBER: 1487-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 2694 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-09-275-265-23

Query Match
Best Local Similarity: 44.58; Score: 59; ID: 2; Length: 2694;
Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;

QY 1821 VLLPAIKFYVGFNFNMHMFNSIQ 1849
DE 1745 VIAVAVQTNWQKLTFWAKHNRFSGIQ 1773

Query Match
Best Local Similarity: 44.58; Score: 59; ID: 2; Length: 2694;
Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;

QY 1821 VLLPAIKFYVGFNFNMHMFNSIQ 1849
DE 1745 VIAVAVQTNWQKLTFWAKHNRFSGIQ 1773

RESULT 64
US-09-275-265-23
Sequence 23, Application US/09275265
Patent No. 6,297,761
GENERAL INFORMATION:
APPLICANT: DELVEY, ROBERT J.
APPLICANT: POLLET, DIRK
APPLICANT: MAERTENS, GEERT
TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
TITLE OF INVENTION: ANTIBODIES TO HEPATITIS C VIRUS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: RIXER & VANERHUYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.40
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/275,265
FILING DATE: 14-DEC-1999
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/98/091,671
FILING DATE: 21-FEB-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/97/420,286
FILING DATE: 14-DEC-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: WO 97/0791/02409
FILING DATE: 13-DEC-1991
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: EP 96134241.2
FILING DATE: 14-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: SADDFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/BOOKLET NUMBER: 1487-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 2694 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-09-275-265-23

Query Match
Best Local Similarity: 44.58; Score: 59; ID: 2; Length: 2694;
Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;

QY 1821 VLLPAIKFYVGFNFNMHMFNSIQ 1849
DE 1745 VIAVAVQTNWQKLTFWAKHNRFSGIQ 1773

RESULT 65
US-08-443-260-4
Sequence 3, Application US/08443260
Patent No. 5942244
GENERAL INFORMATION:
APPLICANT: RALSTON, ROBERT O.
APPLICANT: MARCUS, FRANK
APPLICANT: THEDUM, KENT B.
APPLICANT: GERVAISE, BARBARA A.
APPLICANT: HALL, JOHN A.
TITLE OF INVENTION: HEPATITIS C VIRUS ASIALOGLYCOPROTEINS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESS: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: California
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 08/08443260
FILING DATE: 17-MAY-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: HARRIN, ALISA A.
REGISTRATION NUMBER: 33,895
REFERENCE/CHECK NUMBER: 0154 006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2708
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2955 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Modified-site
LOCATION: 9
OTHER INFORMATION: /note- "There is a heterogeneity at this location; Xaa - Arg or Lys"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 11
OTHER INFORMATION: /note- "There is a heterogeneity at this location; Xaa - Asn or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 176
OTHER INFORMATION: /note- "There is a heterogeneity at this location; Xaa - Ile or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 444
OTHER INFORMATION: /note- "There is a heterogeneity at this location; Xaa - Met or Val"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 603
OTHER INFORMATION: /note- "There is a heterogeneity at this location; Xaa - Ile or Leu"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 848
OTHER INFORMATION: /note- "There is a heterogeneity at this location; Xaa - Asn or Tyr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1114
OTHER INFORMATION: /note- "There is a heterogeneity at this location; Xaa - Pro or Ser"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1117
OTHER INFORMATION: /note- "There is a heterogeneity at this location; Xaa - Ser or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1276
OTHER INFORMATION: /note- "There is a heterogeneity at this location; Xaa - Leu or Pro"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1454
OTHER INFORMATION: /note- "There is a heterogeneity at this location; Xaa - Cys or Tyr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1471
OTHER INFORMATION: /note- "There is a heterogeneity at this location; Xaa - Ser or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1877
OTHER INFORMATION: /note- "There is a heterogeneity at this location; Xaa - Glu or Gly"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1948
OTHER INFORMATION: /note- "There is a heterogeneity at this location; Xaa - His or Leu"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1949
OTHER INFORMATION: /note- "There is a heterogeneity at this location; Xaa - Cys or Ser"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2021
OTHER INFORMATION: /note- "There is a heterogeneity at this location; Xaa - Gly or Val"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2349
OTHER INFORMATION: /note- "There is a heterogeneity at this location; Xaa - Ser or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2385
OTHER INFORMATION: /note- "There is a heterogeneity at this location; Xaa - Phe or Tyr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2386
OTHER INFORMATION: /note- "There is a heterogeneity at this location; Xaa - Ala or Ser"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2502
OTHER INFORMATION: /note- "There is a heterogeneity at this location; Xaa - Phe or Leu"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2690
OTHER INFORMATION: /note- "There is a heterogeneity at this location; Xaa - Gly or Arg"


```

FEATURE:
NAME/KEY: Modified-site
LOCATION: 2502
OTHER INFORMATION: /note- "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Phe or Leu"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2690
OTHER INFORMATION: /note- "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Gly or Arg"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2921
OTHER INFORMATION: /note- "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Arg or Gly"
US-08-442-805A-3

Query Match 0.5%; Score 59; DB 3; Length 2955;
Best Local Similarity 34.5%; Pled. No. 2,1e+03;
Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1821 VLLPAIKETVKGIEKNKNNMGPFMSILQ 1849
I: I:: :::: I I I I I I I I
Db 1745 VIAPVGTNWKLETFWAKIMNPFISILQ 1773

RESULT 67
US-08-443-900A-3
Sequence 3, Application IIS/08443900A
Patent No. 6074852
GENERAL INFORMATION:
APPLICANT: RALSTON, ROBERT O.
APPLICANT: MARCUS, FRANK
APPLICANT: THUDION, KENI B.
APPLICANT: GERVAISE, BARBARA A.
APPLICANT: HALL, JOHN A.
TITLE OF INVENTION: HEPATITIS C VIRUS ANTALOGYCOPROTEINS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: California
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/443,900A
FILING DATE: 17-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: HARBIN, ALISA A.
REGISTRATION NUMBER: 33,895
REFERENCE/INVENT NUMBER: 0154,004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2708
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2955 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Modified-site
LOCATION: 9
OTHER INFORMATION: /note- "There is a heterogeneity at

```



```

OTHER INFORMATION: Xaa which is either Gly or Val"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2349
OTHER INFORMATION: /note= "A heterogeneity exists at
OTHER INFORMATION: Xaa which is either Thr or Ser"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2385
OTHER INFORMATION: /note= "A heterogeneity exists at
OTHER INFORMATION: Xaa which is either Tyr or Phe"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2386
OTHER INFORMATION: /note= "A heterogeneity exists at
OTHER INFORMATION: Xaa which is either Ser or Ala"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2502
OTHER INFORMATION: /note= "A heterogeneity exists at
OTHER INFORMATION: Xaa which is either Leu or Phe"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2690
OTHER INFORMATION: /note= "A heterogeneity exists at
OTHER INFORMATION: Xaa which is either Arg or Gly"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2921
OTHER INFORMATION: /note= "A heterogeneity exists at
OTHER INFORMATION: Xaa which is either Arg or Gly"
US-08-444-818-124

Query Match 0.5%; Score 59; DB 4; Length 2955;
Best Local Similarity 34.5%; Pred. R6; 2.1e+03;
Matches 16; Conservative 5; Mismatches 16; Indels 0; Gaps 0;

QY 1821 VLLPALKFYKFKYQKRNKNNM:PFMSILQ 1849
1: 11: 1111 1 1 11 1111
DB 1745 VIAPVQTRWQKLETWAKHWHFISILQ 1773

RESULT 69
US-08-249-843-3
Sequence 3, Application US/08249843
Patent No. 6274148
GENERAL INFORMATION:
APPLICANT: RALSTON, ROBERT O.
APPLICANT: MARCUS, FRANK
APPLICANT: THUDION, KENT B.
APPLICANT: GERVAISE, BARBARA A.
APPLICANT: HALL, JOHN A.
TITLE OF INVENTION: HEPATITIS C VIRUS ASIALOGLYCOPROTEINS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: California
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/249,843
FILING DATE: 26-MAY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: HARBIN, ALISA A.

```

```

REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0154 003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2708
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2455 amino acids
TYPE: amino acid
STRANDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Modified-site
LOCATION: 9
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Arg or Lys"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 11
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Asn or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 176
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Ile or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 334
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Met or Val"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 603
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Ile or Leu"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 848
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Asn or Tyr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1114
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Pro or Ser"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1117
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Ser or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1276
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Leu or Pro"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1454
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Cys or Tyr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1471
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Ser or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1877
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Glu or Gly"
FEATURE:

```



```

1 NAME/KEY: Modified site
2 LOCATION: 144
3 OTHER INFORMATION: Note "There is a heterogeneity at
4 THIS INFORMATION: this location; Xaa His or Lys"
5 FEATURE:
6 NAME/KEY: Modified site
7 LOCATION: 144
8 OTHER INFORMATION: Note "There is a heterogeneity at
9 THIS INFORMATION: this location; Xaa Cys or Ser"
10 FEATURE:
11 NAME/KEY: Modified site
12 LOCATION: 2121
13 OTHER INFORMATION: Note "There is a heterogeneity at
14 THIS INFORMATION: this location; Xaa Gly or Val"
15 FEATURE:
16 NAME/KEY: Modified site
17 LOCATION: 244
18 OTHER INFORMATION: Note "There is a heterogeneity at
19 THIS INFORMATION: this location; Xaa Ser or Thr"
20 FEATURE:
21 NAME/KEY: Modified site
22 LOCATION: 2502
23 OTHER INFORMATION: Note "There is a heterogeneity at
24 THIS INFORMATION: Xaa His or Lys"
25 FEATURE:
26 NAME/KEY: Modified site
27 LOCATION: 2509
28 OTHER INFORMATION: Note "There is a heterogeneity at
29 THIS INFORMATION: Xaa Gly or Ala"
30 FEATURE:
31 NAME/KEY: Modified site
32 LOCATION: 2921
33 OTHER INFORMATION: Note "There is a heterogeneity at
34 THIS INFORMATION: this location; Xaa Ala or Gly"
35 US 09 249 640-3

```

```

Query Match
Best Local Similarity 4.58; Score 59; Db 4; Length 2955;
Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

```

```

QY 1821 VILPAIKKIVYTERKNNKNNHPEMSILQ 1849
14 144 VIAFAVQTNKLEIKWAKBNHPTLQ 1774

```

```

RESULT 71
US 09 144 640-3
Sequence 144; Application US/0644140
Patent No. 6,606,847
GENERAL INFORMATION:
APPLICANT: Chiron, David Y.
APPLICANT: Partner, William J.
TITLE OF INVENTION: NARROW GLAUCOMATIVES AND VACCINES
NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 1500 Horton Street
CITY: Emeryville
STATE: CA
ZIP: 94608-2016
COMPUTER READABLE FORM:

```

```

1 MEDIUM TYPE: Floppy disk
2 COMPUTER: IBM PC compatible
3 OPERATING SYSTEM: PC-DOS/MS-DOS
4 SOFTWARE: Patent In Release #1.0, Version #1.0
5 CURRENT APPLICATION DATA:
6 APPLICATION NUMBER: 09/009,444,818
7 FILING DATE:
8 CLASSIFICATION: 424
9 PRIOR APPLICATION DATA:
10 APPLICATION NUMBER: US/08/403,590
11 FILING DATE: 14-MAR-1995
12 ATTORNEY/AGENT INFORMATION:
13 NAME: Barlin, Alison A.
14 REGISTRATION NUMBER: 33,895
15 REFERENCE/KEY NUMBER: 611,902
16 TELECOMMUNICATION INFORMATION:
17 TELEPHONE: (508) 359-3885
18 TELEFAX: (508) 359-3885
19 INFORMATION FOR SEQ ID NO: 148:
20 SEQUENCE CHARACTERISTICS:
21 LENGTH: 2945 amino acids
22 TYPE: amino acid
23 TOPOLOGY: linear
24 MOLECULE TYPE: protein
25 US 08 444-818-148

```

```

Query Match
Best Local Similarity 4.58; Score 59; Db 4; Length 2955;
Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

```

```

QY 1821 VILPAIKKIVYTERKNNKNNHPEMSILQ 1849
14 144 VIAFAVQTNKLEIKWAKBNHPTLQ 1774

```

```

RESULT 71
US 08 453-552-2
Sequence 2; Application US/0845352
Patent No. 5667992
GENERAL INFORMATION:
APPLICANT: CASEY, JAMES M.
APPLICANT: BOEL, SUZANNE L.
APPLICANT: ZETK, BILLY J.
APPLICANT: YAMAGUCHI, JULIE
APPLICANT: FRAIL, DONALD E.
APPLICANT: DESAI, SURESH M.
APPLICANT: DEVARE, SUSHIL G.
TITLE OF INVENTION: MAMMALIAN EXPRESSION SYSTEMS FOR HCV
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES
STREET: ONE ABBOTT PARK ROAD
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 09/009,453,552
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: POREMSKI, PRISCILLA E.
REGISTRATION NUMBER: 33,287
REFERENCE/KEY NUMBER: 5131.US.11
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365

```

DB 1745 VIAPAVQTNWQKLETFWAKHWMNFISGIQ 1773

RESULT 73

US-08-440-542-36

; Sequence 36, Application US/08440542

; Patent No. 5670153

; GENERAL INFORMATION:

; APPLICANT: Weiner, Amy J.

; APPLICANT: Houghton, Michael

; TITLE OF INVENTION: Immunoreactive Polypeptide Compositions

; NUMBER OF SEQUENCES: 45

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Chiron Corporation

; STREET: 4560 Horton Street

; CITY: Emeryville

; STATE: CA

; COUNTRY: USA

; ZIP: 94608

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/440-542

; FILING DATE: 12-MAY-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION NUMBER: US/08/231,368

; FILING DATE:

; APPLICATION NUMBER: US 07/759,575

; FILING DATE: 13-SEP-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: McClung, Barbara G.

; REGISTRATION NUMBER: 33,113

; REFERENCE/DOCKET NUMBER: 0205,001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (510) 601-2708

; TELEFAX: (510) 655-3542

; INFORMATION FOR SEQ ID NO: 36:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3011 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-440-542-36

Query Match 0.5%; Score 59; DB 1; Length 3011;

Best Local Similarity 34.5%; Pred. No. 2.2e+03;

Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1821 VLIPAIKTKYKQIEKNKNNHMGPMSTIIQ 1849

1: 111: 111: 11 11 11 11

Db 1745 VIAPAVQTNWQKLETFWAKHWMNFISGIQ 1773

RESULT 74

US-07-910-760-10

; Sequence 10, Application US/07910760

; Patent No. 5683864

; GENERAL INFORMATION:

; APPLICANT: Houghton, Michael

; APPLICANT: Choo, Qui-lim

; TITLE OF INVENTION: Combinations of Hepatitis B virus (HBV)

; TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Chiron Corporation

; STREET: P O Box 8047 (Int. Prop. R-440)

QY 1821 VLIPAIKTKYKQIEKNKNNHMGPMSTIIQ 1849

1: 111: 111: 11 11 11 11

Db 1745 VIAPAVQTNWQKLETFWAKHWMNFISGIQ 1773

RESULT 72

US-08-440-104-36

; Sequence 36, Application US/08440103

; Patent No. 5670152

; GENERAL INFORMATION:

; APPLICANT: Weiner, Amy J.

; APPLICANT: Houghton, Michael

; TITLE OF INVENTION: Immunoreactive Polypeptide Compositions

; NUMBER OF SEQUENCES: 45

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Chiron Corporation

; STREET: 4560 Horton Street

; CITY: Emeryville

; STATE: CA

; COUNTRY: USA

; ZIP: 94608

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/440,103

; FILING DATE: 12-MAY-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION NUMBER: US/08/231,368

; FILING DATE:

; APPLICATION NUMBER: US 07/759,575

; FILING DATE: 13-SEP-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: McClung, Barbara G.

; REGISTRATION NUMBER: 33,113

; REFERENCE/DOCKET NUMBER: 0205,001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (510) 601-2708

; TELEFAX: (510) 655-3542

; INFORMATION FOR SEQ ID NO: 36:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3011 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-440-104-36

Query Match 0.5%; Score 59; DB 1; Length 3011;

Best Local Similarity 34.5%; Pred. No. 2.2e+03;

Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1821 VLIPAIKTKYKQIEKNKNNHMGPMSTIIQ 1849

1: 111: 111: 11 11 11 11

CITY: Emeryville
 STATE: CA
 COUNTRY: U.S.A.
 ZIP: 94602-6007
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 MEDIUM: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/97/410,760
 FILING DATE: 12 MAY 1997
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: BUCKLE & SUTHERLAND, Robert P.
 REFERENCE NUMBER: 13417
 TELEPHONE: (510) 601-1942
 TELEFAX: (510) 601-2702
 INVENTOR INFORMATION:
 NAME: EMERYVILLE, CA, USA
 LENGTH: 401 amino acids
 TYPE: amine acid
 TOPOLOGY: Linear
 MOLECULE TYPE: protein
 US-09-603-665-5

Query Match: 0.58; Score 59; DB 1; Length 401;
 Best Local Similarity: 34.58; Pct. No. 2,200,000;
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;

27 1821 VLFAVATNKVLEWENMHPFNSIQ 1849
 18 1745 VIAFAVQTNWQLETFWAKHWNFISQIQ 1774

RESULT 76
 US-08-241-368-46
 Sequence 36, Application US/08241368
 Patent No. 5756412
 GENERAL INFORMATION:
 APPLICANT: Weinert, Amy J.
 APPLICANT: Houghton, Michael
 TITLE OF INVENTION: Immunoreactive polypeptide compositions
 NUMBER OF SEQUENCES: 45
 CORRESPONDENCE ADDRESS:
 ADDRESS: Chiron Corporation
 STREET: 4560 Horton Street
 CITY: Emeryville
 STATE: CA
 COUNTRY: USA
 ZIP: 94608
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/97/410,760
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/97/410,760
 FILING DATE: 13-SEP-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: McClum, Barbara G.
 REGISTRATION NUMBER: 33,111
 REFERENCE NUMBER: 0205,001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (510) 601-2708
 TELEFAX: (510) 655-3542
 INFORMATION FOR SEQ ID NO: 46:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 401 amino acids
 TYPE: amine acid
 TOPOLOGY: single
 CLANDBLASS: linear
 MOLECULE TYPE: protein
 US-08-241-368-46

Query Match: 0.58; Score 59; DB 1; Length 401;
 Best Local Similarity: 34.58; Pct. No. 2,200,000;
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;

27 1821 VLFAVATNKVLEWENMHPFNSIQ 1849
 18 1745 VIAFAVQTNWQLETFWAKHWNFISQIQ 1774

CITY: Emeryville
 STATE: CA
 COUNTRY: U.S.A.
 ZIP: 94602-6007
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 MEDIUM: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/97/410,760
 FILING DATE: 12 MAY 1997
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: BUCKLE & SUTHERLAND, Robert P.
 REFERENCE NUMBER: 13417
 TELEPHONE: (510) 601-1942
 TELEFAX: (510) 601-2702
 INVENTOR INFORMATION:
 NAME: EMERYVILLE, CA, USA
 LENGTH: 401 amino acids
 TYPE: amine acid
 TOPOLOGY: Linear
 MOLECULE TYPE: protein
 US-08-440-519-10

Query Match: 0.58; Score 59; DB 1; Length 401;
 Best Local Similarity: 34.58; Pct. No. 2,200,000;
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;

27 1821 VLFAVATNKVLEWENMHPFNSIQ 1849
 18 1745 VIAFAVQTNWQLETFWAKHWNFISQIQ 1774

RESULT 76
 US-08-241-368-46
 Sequence 36, Application US/08241368
 Patent No. 5756412
 GENERAL INFORMATION:
 APPLICANT: Weinert, Amy J.
 APPLICANT: Houghton, Michael
 TITLE OF INVENTION: Immunoreactive polypeptide compositions
 NUMBER OF SEQUENCES: 45
 CORRESPONDENCE ADDRESS:
 ADDRESS: Chiron Corporation
 STREET: 4560 Horton Street
 CITY: Emeryville
 STATE: CA
 COUNTRY: USA
 ZIP: 94608
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/97/410,760
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/97/410,760
 FILING DATE: 13-SEP-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: McClum, Barbara G.
 REGISTRATION NUMBER: 33,111
 REFERENCE NUMBER: 0205,001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (510) 601-2708
 TELEFAX: (510) 655-3542
 INFORMATION FOR SEQ ID NO: 46:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 401 amino acids
 TYPE: amine acid
 TOPOLOGY: single
 CLANDBLASS: linear
 MOLECULE TYPE: protein
 US-08-241-368-46

Query Match: 0.58; Score 59; DB 1; Length 401;
 Best Local Similarity: 34.58; Pct. No. 2,200,000;
 Matches: 10; Conservative: 9; Mismatches: 10; Indels: 0; Gaps: 0;

27 1821 VLFAVATNKVLEWENMHPFNSIQ 1849
 18 1745 VIAFAVQTNWQLETFWAKHWNFISQIQ 1774

us-09-603-665-5.ra1

Fri Nov 9 15:42:38 2001

ADDRESSES: ABBOTT LABORATORIES D377/AP6D
 STREET: ONE ABBOTT PARK ROAD
 CITY: ABBOTT PARK
 STATE: IL
 COUNTRY: USA
 ZIP: 60064-3500
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/710,637
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/144,099
 FILING DATE:
 APPLICATION NUMBER: US/07/830,024
 FILING DATE: 01-JAN-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: POREMBSKI, PRISCILLA E.
 REGISTRATION NUMBER: 33,207
 REFERENCE/DOCKET NUMBER: 5131.US.01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 708-937-6365
 TELEFAX: 708-937-9556
 INFORMATION FOR SEQ ID NO. 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3011 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-710-637-2

Query Match 0.5%; Score 59; DB 2; Length 3011;
 Best Local Similarity 34.5%; Pred. No. 2.2e+03;
 Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1821 VLLIATKTYKQIFKKNWNNMCPMSIIQ 1849
 Db 1745 VIAPAVQTNWQLETFWAKHWNFISGIQ 1773

RESULT 79
 US-08-833-678A-6
 Sequence 6, Application US/08833678A
 Patent No. 5989905
 GENERAL INFORMATION:
 APPLICANT: HOUGHTON, MICHAEL
 APPLICANT: CHOO, GUL-LIM
 APPLICANT: HAN, JANG
 APPLICANT: CHOP, JOONHO
 TITLE OF INVENTION: HCV NS3 PROTEIN FRAGMENTS HAVING
 TITLE OF INVENTION: HELICASE ACTIVITY AND IMPROVED SOLUBILITY
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: CHIRON CORPORATION
 STREET: Intellectual Property - R440, P.O. Box 8097
 CITY: Emeryville
 STATE: California
 COUNTRY: U.S.A.
 ZIP: 94662-8097
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/833,678A
 FILING DATE:

RESULT 77
 US-08-440-210-36
 Sequence 46, Application US/08440210
 Patent No. 5765845
 GENERAL INFORMATION:
 APPLICANT: Weiner, Amy J.
 APPLICANT: Houghton, Michael
 TITLE OF INVENTION: Immunoreactive Polypeptide Compositions
 NUMBER OF SEQUENCES: 45
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Chiron Corporation
 STREET: 4560 Horton Street
 CITY: Emeryville
 STATE: CA
 COUNTRY: USA
 ZIP: 94608
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/440,210
 FILING DATE: 12-MAY-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/144,099
 FILING DATE:
 APPLICATION NUMBER: US/07/759,575
 FILING DATE:
 APPLICATION NUMBER: 13-SEP-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: McClung, Barbara G.
 REGISTRATION NUMBER: 33,113
 REFERENCE/DOCKET NUMBER: 0205.001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (510) 601-2708
 TELEFAX: (510) 655-3542
 INFORMATION FOR SEQ ID NO. 36:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3011 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-440-210-36

Query Match 0.5%; Score 59; DB 1; Length 3011;
 Best Local Similarity 34.5%; Pred. No. 2.2e+03;
 Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1821 VLLIATKTYKQIFKKNWNNMCPMSIIQ 1849
 Db 1745 VIAPAVQTNWQLETFWAKHWNFISGIQ 1773

RESULT 78
 US-08-710-637-2
 Sequence 2, Application US/08710637
 Patent No. 5854001
 GENERAL INFORMATION:
 APPLICANT: CASEY, JAMES M.
 APPLICANT: BEDE, SUZANNE L.
 APPLICANT: ZECK, RILEY J.
 APPLICANT: YAMAGUCHI, JULIE
 APPLICANT: FRAIL, DONALD F.
 APPLICANT: DESAI, SURESH M.
 APPLICANT: DEVARE, SUSHIL G.
 TITLE OF INVENTION: MAMMALIAN EXPRESSION SYSTEMS FOR HCV
 TITLE OF INVENTION: PROTEINS
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:

CLASSIFICATION:
 PUBLICATION NUMBER: US 00/000,000
 FILING DATE: 15 SEP 1995
 ATTORNEY: J. J. J. J. J.
 NAME: J. J. J. J. J.
 PUBLICATION NUMBER: 33,885
 REFERENCE NUMBER: 0100,005
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (510) 555-4542
 TELEFAX: (510) 555-4542
 INDEX: 0/0
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4011 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 FEATURE:
 NAME/KEY: duplication
 LOCATION: 9
 OTHER INFORMATION: Znote "There exists a
 OTHER INFORMATION: heterogeneity at this position - Xaa - Cys or Arg"
 FEATURE:
 NAME/KEY: duplication
 LOCATION: 11
 OTHER INFORMATION: Znote "There exists a
 OTHER INFORMATION: heterogeneity at this position - Xaa - Asn or Thr"
 FEATURE:
 NAME/KEY: duplication
 LOCATION: 176
 OTHER INFORMATION: Znote "There exists a
 OTHER INFORMATION: heterogeneity at this position - Xaa - Ile or Thr"
 FEATURE:
 NAME/KEY: duplication
 LOCATION: 44
 OTHER INFORMATION: Znote "There exists a
 OTHER INFORMATION: heterogeneity at this position - Xaa - Met or Val"
 FEATURE:
 NAME/KEY: duplication
 LOCATION: 848
 OTHER INFORMATION: Znote "There exists a
 OTHER INFORMATION: heterogeneity at this position - Xaa - Leu or Ile"
 FEATURE:
 NAME/KEY: duplication
 LOCATION: 1104
 OTHER INFORMATION: Znote "There exists a
 OTHER INFORMATION: heterogeneity at this position - Xaa - Tyr or Asn"
 FEATURE:
 NAME/KEY: duplication
 LOCATION: 1276
 OTHER INFORMATION: Znote "There exists a
 OTHER INFORMATION: heterogeneity at this position - Xaa - Pro or Ser"
 FEATURE:
 NAME/KEY: duplication
 LOCATION: 1454
 OTHER INFORMATION: Znote "There exists a
 OTHER INFORMATION: heterogeneity at this position - Xaa - Cys or Tyr"
 FEATURE:
 NAME/KEY: duplication
 LOCATION: 1771
 OTHER INFORMATION: Znote "There exists a

OTHER INFORMATION: heterogeneity at this position - Xaa - Thr or Ser"
 FEATURE:
 NAME/KEY: duplication
 LOCATION: 1877
 OTHER INFORMATION: Znote "There exists a
 OTHER INFORMATION: heterogeneity at this position - Xaa - Glu or Gly"
 FEATURE:
 NAME/KEY: duplication
 LOCATION: 1948
 OTHER INFORMATION: Znote "There exists a
 OTHER INFORMATION: heterogeneity at this position - Xaa - Leu or His"
 FEATURE:
 NAME/KEY: duplication
 LOCATION: 1949
 OTHER INFORMATION: Znote "There exists a
 OTHER INFORMATION: heterogeneity at this position - Xaa - Ser or Cys"
 FEATURE:
 NAME/KEY: duplication
 LOCATION: 2021
 OTHER INFORMATION: Znote "There exists a
 OTHER INFORMATION: heterogeneity at this position - Xaa - Gly or Val"
 FEATURE:
 NAME/KEY: duplication
 LOCATION: 2349
 OTHER INFORMATION: Znote "There exists a
 OTHER INFORMATION: heterogeneity at this position - Xaa - Tyr or Phe"
 FEATURE:
 NAME/KEY: duplication
 LOCATION: 2485
 OTHER INFORMATION: Znote "There exists a
 OTHER INFORMATION: heterogeneity at this position - Xaa - Tyr or Phe"
 FEATURE:
 NAME/KEY: duplication
 LOCATION: 2486
 OTHER INFORMATION: Znote "There exists a
 OTHER INFORMATION: heterogeneity at this position - Xaa - Ser or Ala"
 FEATURE:
 NAME/KEY: duplication
 LOCATION: 2502
 OTHER INFORMATION: Znote "There exists a
 OTHER INFORMATION: heterogeneity at this position - Xaa - Leu or Phe"
 FEATURE:
 NAME/KEY: duplication
 LOCATION: 2690
 OTHER INFORMATION: Znote "There exists a
 OTHER INFORMATION: heterogeneity at this position - Xaa - Arg or Gly"
 FEATURE:
 NAME/KEY: duplication
 LOCATION: 2921
 OTHER INFORMATION: Znote "There exists a
 OTHER INFORMATION: heterogeneity at this position - Xaa - Leu or Pro"
 FEATURE:
 NAME/KEY: duplication
 LOCATION: 2996
 OTHER INFORMATION: Znote "There exists a
 OTHER INFORMATION: heterogeneity at this position - Xaa - Leu or Pro"

US-08-833-678A-6

Query Match: 0.58; Score 59; 100 Z; Length 4011;

Best Local Similarity: 44.5%; Pred. No. 2,400,000;

Matches: 10; Conservative: 9; Mismatches: 10; Labels: 0; Gaps: 0;

CY: 1821 VLFAIKKIVQILKAKSKNEMGPPFMSILQ 1849

13 133 3333 11 1 1 1

PE: 1745 VIAPAVQINWQLETFWAKIMMNNF 831Q 1773

RESULT: 80

US-08-444-818-177

Sequence: 177; Applicant: US-08474818

Patent No. 6150087

GENERAL INFORMATION:

APPLICANT: Chiron, David Y.
TITLE OF INVENTION: NANV Diagnostics and Vaccines
NUMBER OF SEQUENCES: 777
CORRESPONDENCE ADDRESS:
ADDRESS: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.40
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-09-603-665-5
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US-09-603-665-5
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Harbin, Alisa A.
REGISTRATION NUMBER: 43,895
REFERENCE/DOCKET NUMBER: 0110.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508) 359-3876
TELEFAX: (508) 359-3885
INFORMATION FOR SEQ ID NO: 177:
SEQUENCE CHARACTERISTICS:
LENGTH: 3011 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Duplication
LOCATION: 9
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa - Lys or Arg"
FEATURE:
NAME/KEY: Duplication
LOCATION: 11
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Asn or Thr"
FEATURE:
NAME/KEY: Duplication
LOCATION: 176
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa - Ile or Thr"
FEATURE:
NAME/KEY: Duplication
LOCATION: 334
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Met or Val"
FEATURE:
NAME/KEY: Duplication
LOCATION: 603
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Leu or Ile"
FEATURE:
NAME/KEY: Duplication
LOCATION: 848
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Tyr or Asn"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1114
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Pro or Ser"
FEATURE:

NAME/KEY: Duplication
LOCATION: 1117
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa - Ser or Thr"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1276
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa - Pro or Leu"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1454
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa - Cys or Tyr"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1471
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa - Thr or Ser"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1877
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa - Glu or Gly"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1948
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Leu or His"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1949
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Ser or Cys"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2021
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa - Gly or Val"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2349
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa - Thr or Ser"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2385
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Tyr or Phe"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2386
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Ser or Ala"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2502
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Leu or Phe"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2690
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Arg or Gly"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2921
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa - Arg or Gly"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2996
OTHER INFORMATION: /note= "There exists a

```

OTHER INFORMATION: heterogeneity at this position - Xaa - Leu or Pro"
US 09 444 616 177
FEATURE:
NAME/KEY: Duplication
LOCATION: 603
OTHER INFORMATION: heterogeneity at this position - Xaa - Met or Val"
NAME/KEY: Duplication
LOCATION: 603
OTHER INFORMATION: heterogeneity at this position - Xaa - Leu or Ile"
FEATURE:
NAME/KEY: Duplication
LOCATION: 848
OTHER INFORMATION: heterogeneity at this position - Xaa - Tyr or Asn"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1114
OTHER INFORMATION: heterogeneity at this position - Xaa - Pro or Ser"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1117
OTHER INFORMATION: heterogeneity at this position - Xaa - Ser or Thr"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1276
OTHER INFORMATION: heterogeneity at this position - Xaa - Pro or Leu"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1454
OTHER INFORMATION: heterogeneity at this position - Xaa - Cys or Tyr"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1471
OTHER INFORMATION: heterogeneity at this position - Xaa - Thr or Ser"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1877
OTHER INFORMATION: heterogeneity at this position - Xaa - Glu or Gly"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1948
OTHER INFORMATION: heterogeneity at this position - Xaa - Leu or His"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1949
OTHER INFORMATION: heterogeneity at this position - Xaa - Ser or Cys"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2021
OTHER INFORMATION: heterogeneity at this position - Xaa - Gly or Val"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2449
OTHER INFORMATION: heterogeneity at this position - Xaa - Thr or Ser"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2485
OTHER INFORMATION: heterogeneity at this position - Xaa - Tyr or Thr"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2486
OTHER INFORMATION: heterogeneity at this position - Xaa - Ser or Ala"

```


PCT-US93-00907-2

Sequence 2, Application PC/PUS9300907
 GENERAL INFORMATION:
 APPLICANT: CASEY, JAMES M.
 APPLICANT: BOBE, SUZANNE L.
 APPLICANT: ZECK, BILLY J.
 APPLICANT: YAMAGUCHI, JULIE
 APPLICANT: FRAIL, DONALD E.
 APPLICANT: DESAI, SURESH M.
 APPLICANT: DEVAIRE, SUSHIL G.
 TITLE OF INVENTION: MAMMALIAN EXPRESSION CYCHEME FOR HCV
 TITLE OF INVENTION: PROTEINS
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: ABBOTT LABORATORIES 3477 AVE C
 STREET: ONE ABBOTT PARK ROAD
 CITY: ABBOTT PARK
 STATE: IL
 COUNTRY: USA
 ZIP: 60064-1500

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US93/00907
 FILING DATE: 19930129

CLASSIFICATION:

AUTORNEY/AGENT INFORMATION:
 NAME: PEREMBSKI, PRISCILLA E.
 REGISTRATION NUMBER: 33,207
 REFERENCE/DOCKET NUMBER: 5131.PC.01

TELECOMMUNICATION INFORMATION:

TELEPHONE: 708-937-6365
 TELEFAX: 708-937-9556

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 3011 amino acids

TYPE: AMINO ACID

STRANDEDNESS: Single

TOPOLOGY: Linear

MOLECULE TYPE: protein

PCT-US93-00907-2

Query Match

Best Local Similarity 6.5%; Score 59; DB 5; Length 3611;

Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 182: VLLTAALQKRVWTFLEHISVLEKTIISGASQANIRLETKKTLAT 1815

DB 1745: VLLTAALQKRVWTFLEHISVLEKTIISGASQANIRLETKKTLAT 1815

RESULT 86

US-08-209-182C-4

Sequence 4, Application US/08209182C

Patent No. 5545537

GENERAL INFORMATION:

APPLICANT: Skelly, Susan M.

APPLICANT: Snowdard, John R.

APPLICANT: Fowlkes, Dana M.

TITLE OF INVENTION: Cysteine Depleted IL-6 Muteins

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Imclone Systems Incorporated

STREET: 180 Varick Street

CITY: New York

STATE: New York

COUNTRY: United States

ZIP: 10014

Query Match

Best Local Similarity 44.5%; Score 59; DB 5; Length 3611;

Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 182: VLLTAALQKRVWTFLEHISVLEKTIISGASQANIRLETKKTLAT 1815

DB 1745: VLLTAALQKRVWTFLEHISVLEKTIISGASQANIRLETKKTLAT 1815

RESULT 87

PCT-US92-05612-4

Sequence 4, Application PC/PUS9205612

GENERAL INFORMATION:

APPLICANT: Skelly, Susan M.

APPLICANT: Snowdard, John N.

APPLICANT: Fowlkes, Dana M.

TITLE OF INVENTION: Cysteine Depleted IL-6 Muteins

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Imclone Systems Incorporated

STREET: 180 Varick Street

CITY: New York

STATE: New York

COUNTRY: United States

ZIP: 10014

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: 95/08/209,182C

FILING DATE: 10-MAR-1994

PRIOR APPLICATION DATA: US 07/724,698

APPLICATION NUMBER: 02-101-1991

FILING DATE: 02-101-1991

AUTORNEY/AGENT INFORMATION:

NAME: Feit, Irving N.

REGISTRATION NUMBER: 28,601

REFERENCE/DOCKET NUMBER: SKE-1-PD

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-645-1405

TELEFAX: 212-645-2054

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 181 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-209-182C-4

Query Match

Best Local Similarity 22.5%; Score 58; DB 1; Length 181;

Matches 18; Conservative 16; Mismatches 46; Indels 0; Gaps 0;

QY 1756: VLLTAALQKRVWTFLEHISVLEKTIISGASQANIRLETKKTLAT 1815

DB 28: VLLTAALQKRVWTFLEHISVLEKTIISGASQANIRLETKKTLAT 1815

QY 1816: TLAPVLLPAIKTKYQIEK 1835

DB 88: TLAPVLLPAIKTKYQIEK 1835

RESULT 87

PCT-US92-05612-4

Sequence 4, Application PC/PUS9205612

GENERAL INFORMATION:

APPLICANT: Skelly, Susan M.

APPLICANT: Snowdard, John N.

APPLICANT: Fowlkes, Dana M.

TITLE OF INVENTION: Cysteine Depleted IL-6 Muteins

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Imclone Systems Incorporated

STREET: 180 Varick Street

CITY: New York

STATE: New York

COUNTRY: United States

ZIP: 10014

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US92/05612

FILING DATE: 19920702

CLASSIFICATION: 530

AUTORNEY/AGENT INFORMATION:

NAME: Feit, Irving N.

REGISTRATION NUMBER: 28,601

REFERENCE/DOCKET NUMBER: SKE-1-PT

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-645-1405

TELEFAX: 212-645-2054

COMPUTER: IBM Compatible
 OPERATING SYSTEM: IBM P.C. DOS 5.0
 SOFTWARE: Word Perfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/357,642A
 FILING DATE: December 15, 1994
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Warburg, Richard J.
 REGISTRATION NUMBER: 32,327
 REFERENCE/DOCKET NUMBER: 32,327
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 LENGTH: 1009
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-357-642A-1

Query Match 0.58; Score 58; DB 2; Length 1009;
 Best Local Similarity 23.3%; Pred. No. 5.6e+02;
 Matches 10; Conservative 13; Mismatches 20; Indels 0; Gaps 0;
 QY 1203 YWQEVTLLELLQHKRLSPQIIIVPTIFNIUSCTEPDQEQ 1245
 DB 627 FLEKNKDVICVLEKGDPLKPDLCPPVITYIMFCWYDPSDR 669

RESULT 96
 US-08-460-626-1
 Sequence 1; Application US/08460626
 Patent No. 5937815
 GENERAL INFORMATION:
 APPLICANT: SIMA LEV
 TITLE OF INVENTION: PVA 2 RELATED PROJECTS ARE
 TITLE OF INVENTION: METHODS
 NUMBER OF SEQUENCES: 25
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: 633 West Fifth Street
 CITY: Suite 4700
 CITY: Los Angeles
 STATE: California
 COUNTRY: U.S.A.
 ZIP: 90071-2066
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 OPERATING SYSTEM: storage
 SOFTWARE: IBM Compatible
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/460,626
 FILING DATE: June 2, 1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/357,642
 FILING DATE: December 15, 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Warburg, Richard J.
 REGISTRATION NUMBER: 32,327
 REFERENCE/DOCKET NUMBER: 3211/121
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 LENGTH: 1009
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-460-626-1

Query Match 0.58; Score 58; DB 2; Length 1009;
 Best Local Similarity 23.3%; Pred. No. 5.6e+02;
 Matches 10; Conservative 13; Mismatches 20; Indels 0; Gaps 0;
 QY 1203 YWQEVTLLELLQHKRLSPQIIIVPTIFNIUSCTEPDQEQ 1245
 DB 627 FLEKNKDVICVLEKGDPLKPDLCPPVITYIMFCWYDPSDR 669

RESULT 97
 US-08-477-451-11
 Sequence 11; Application US/08477451
 Patent No. 5928865
 GENERAL INFORMATION:
 APPLICANT: Covacci, Antonello
 TITLE OF INVENTION: Bacteriophage T4
 NUMBER OF SEQUENCES: 46
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Chiron Corporation
 STREET: 4560 Horton Street
 CITY: Emeryville
 STATE: CA
 COUNTRY: USA
 ZIP: 94608-2916
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent in Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/477,451
 FILING DATE: 07-MAR-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: McClung, Barbara G.
 REGISTRATION NUMBER: 33,113
 REFERENCE/DOCKET NUMBER: 0335,002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 510-601-2708
 TELEFAX: 510-655-3542
 INFORMATION FOR SEQ ID NO. 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1781 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-477-451-11

Query Match 0.58; Score 58; DB 2; Length 1781;
 Best Local Similarity 38.5%; Pred. No. 1.3e+03;
 Matches 10; Conservative 6; Mismatches 10; Indels 0; Gaps 0;
 QY 1633 WKFTIVTPEFLKIVPTIATVQPKRKE 1658
 DB 547 WSKKITTKILQASLAIITQKMKO 572
 RESULT 98

```

Query Match      0.5%; Score 58; DB 3; Length 3011;
Best Local Similarity 34.5%; Pred. NO. 3e+04;
Matches 10; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1821 VLLPAIKETKYGQTEKKNNKNNHNGSPNSILQ 1849
      | | | | | | | | | | | | | | | |
IB 1745 VITPAVSTNCKLEIVFAKIMWRRFNSGLQ 1773

RESULT 100
OS-09-014-416-1
; Sequence 1, Application OS/00014216
; Patient No. 6153421

```

```

: APPLICANT: Yanagi, Masayuki
: APPLICANT: Buki, Jens
: APPLICANT: Emerson, Susanne U.
: APPLICANT: Parell, Robert H.
: TITLE OF INVENTION: CLONED GENES OF INFECTIOUS RELAVITIS C VIRUSES AND
: FIELD OF INVENTION: USES THEREOF
: FILE REFERENCE: 20264276
: CURRENT APPLICATION NUMBER: US09/014,416
: CURRENT FILING DATE: 1998-01-27
: EARLIER APPLICATION NUMBER: US 09/054,062
: EARLIER FILING DATE: 1997-07-18
: NUMBER OF SEQ. ID NOS: 65

```

; SEQ ID NO 1
 ; LENGTH: 3011
 ; TYPE: PRT
 ; ORGANISM: Hepatitis C virus
 OS=09-014-416-1

Best Local Similarity 34.5%; Score 58; DB 4; Length 3011;
Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1821 VLLPAIKKTYKQIEKNMKNHMGPPMSILQ 1849
I: I I I I I I I I I I I I I I I I
Db 1745 VITPAVQTNWQKLEFVWAKHMMNFISGIQ 1773

RESULT 101

US-09-014-416-5

Sequence 5, Application US/99014476

Patent No. 6153421

GENERAL INFORMATION:

APPLICANT: Yanagii, Masayuki

APPLICANT: Bokh, Jens

APPLICANT: Emerson, Susanne H.

APPLICANT: Farcell, Robert H.

TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND

TITLE OF INVENTION: USES THEREOF

FILE REFERENCE: 20244274

CURRENT APPLICATION NUMBER: 09/09/014,416

CURRENT FILING DATE: 1998-01-27

EARLIER APPLICATION NUMBER: US 60/054,062

EARLIER FILING DATE: 1997-07-18

NUMBER OF SEQ ID NOS: 65

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 5

LENGTH: 3011

TYPE: PKT

ORGANISM: Hepatitis C virus

US-09-014-416-5

Query Match 0.5%; Score 58; DB 4; Length 3011;
Best Local Similarity 34.5%; Score 58; DB 4; Length 3011;
Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1821 VLLPAIKKTYKQIEKNMKNHMGPPMSILQ 1849
I: I I I I I I I I I I I I I I I I
Db 1745 VITPAVQTNWQKLEFVWAKHMMNFISGIQ 1773

RESULT 102

US-09-011-566-2

Sequence 2, Application US/99011566

Patent No. 6127116

GENERAL INFORMATION:

APPLICANT: Rice, Charles, et al.

TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C

TITLE OF INVENTION: VIRUS (HCV) AND USES THEREOF

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESSES:

ADDRESSEE: David A. Jackson, Esq.

STREET: 411 Hackensack Ave., Continental Plaza, 4th

STREET: Floor

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/99/011,566

FILING DATE: 03-MAR-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Jackson, Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 1113-1-006

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3012 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEetical: NO
FRAGMENT TYPE: N terminal
US-08-811-566-2

Query Match 0.5%; Score 58; DB 3; Length 3012;
Best Local Similarity 34.5%; Score 58; DB 3; Length 3012;
Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1821 VLLPAIKKTYKQIEKNMKNHMGPPMSILQ 1849
I: I I I I I I I I I I I I I I I I
Db 1745 VITPAVQTNWQKLEFVWAKHMMNFISGIQ 1773

RESULT 103

US-08-209-1820-8

Sequence 8, Application US/98209182C

Patent No. 5545537

GENERAL INFORMATION:

APPLICANT: Skelly, Susan M.

APPLICANT: Tackney, Charles T.

APPLICANT: Snowwaert, John N.

APPLICANT: Fowlkes, Dana M.

TITLE OF INVENTION: Cysteine Depleted IL 6 Mutains

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESSES:

ADDRESSEE: MacClone Systems Incorporated

STREET: 180 Varick Street

CITY: New York

STATE: New York

COUNTRY: United States

ZIP: 10014

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/98/209,182C

FILING DATE: 10-MAR-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/724,698

FILING DATE: 02-JUL-1991

ATTORNEY/AGENT INFORMATION:

NAME: Pelt, Irving N.

REGISTRATION NUMBER: 28,601

REFERENCE/DOCKET NUMBER: SKE-1-PD

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-645-1405

TELEFAX: 212-645-2054

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 200 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-209-1820-8

Query Match 0.5%; Score 57; DB 1; Length 2007;
Best Local Similarity 32.5%; Score 57; DB 1; Length 2007;
Matches 18; Conservative 16; Mismatches 46; Indels 0; Gaps 0;

QY 1756 YLLSALAALQKVVETLPHFTSPYIFGIIISQVHLEKTISEMSSASQANIRITISIKKTIAT 1815

175 EDVQHQILATSSNQRYVRMLDLQIKSAIKETIKETIVSFQPTQTNQNLQKQPE 222

Query Match 0.5%; Score 57; DB 1; Length 351;
Best Local Similarity 22.4%; Pred. No. 1.5e+02;
Matches 13; Conservative 12; Mismatches 27; Indels 0; Gaps 0;

1779 ETIKAVSNKSSQSHMLGVNPFVHTSQIHPKPLSVSPMSQLSSNPLKKVVE 1536

175 EDVQHQILATSSNQRYVRMLDLQIKSAIKETIKETIVSFQPTQTNQNLQKQPE 222

RESULT 112

US-08-995-654-2
Sequence 2, Application US/08995654
Patent No. 6025138
GENERAL INFORMATION:
APPLICANT: Hawkins, Phillip R.
APPLICANT: Wilde, Craig
APPLICANT: Seilhamer, Jeffrey
TITLE OF INVENTION: HYALURONAN RECEPTOR EXPRESSED IN HUMAN
TITLE OF INVENTION: UMBILICAL VEIN ENDOTHELIAL CELLS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 4174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
COMPUTER: IBM compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/995,654
FILING DATE: December 22, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/700,178
FILING DATE: August 20, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/402,217
FILING DATE: March 10, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PP-0028-2 DIV
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 351 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-995-654-2

Query Match 0.5%; Score 57; DB 3; Length 351;
Best Local Similarity 22.4%; Pred. No. 1.5e+02;
Matches 13; Conservative 18; Mismatches 27; Indels 0; Gaps 0;

1479 ETIPNAVSINSESGFTMGVFNVHTSKQLKPLKPLVSVSSQLSSNPLKKVVE 1536

175 EDVQHQILATSSNQRYVRMLDLQIKSAIKETIKETIVSFQPTQTNQNLQKQPE 222

Query Match 0.5%; Score 57; DB 1; Length 351;
Best Local Similarity 22.4%; Pred. No. 1.5e+02;
Matches 13; Conservative 12; Mismatches 27; Indels 0; Gaps 0;

1779 ETIKAVSNKSSQSHMLGVNPFVHTSQIHPKPLSVSPMSQLSSNPLKKVVE 1536

175 EDVQHQILATSSNQRYVRMLDLQIKSAIKETIKETIVSFQPTQTNQNLQKQPE 222

RESULT 113

US-08-415-823-2
Sequence 2, Application US/08415823
Patent No. 5759538
GENERAL INFORMATION:
APPLICANT: Donovan, William P.
APPLICANT: Tan, Yaping
TITLE OF INVENTION: BACILLUS THURINGIENSIS apt AND apt
TITLE OF INVENTION: PROTEASE GENES AND ALKALINE PROTEASE DEFICIENT AND NEUTRAL
TITLE OF INVENTION: PROTEASE DEFICIENT B.T. STRAINS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Farnitch Schwabe Jacobs & Nadel, P.C.
STREET: 1601 Market Street, 36th floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19103-2398
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/415,823
FILING DATE:
CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 397 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-415-823-2

Query Match 0.5%; Score 57; DB 1; Length 397;
Best Local Similarity 46.7%; Pred. No. 1.9e+02;
Matches 11; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

1720 LTNIMS:PTFM:GVNPMPLDITYSPQVI 1346

175 LSNVQSPHKSVGANVLSKDDKLGFEV 76

RESULT 114

US-09-086-662-2
Sequence 2, Application US/09086662
Patent No. 5962264
GENERAL INFORMATION:
APPLICANT: Donovan, William P.
APPLICANT: Tan, Yaping
TITLE OF INVENTION: BACILLUS THURINGIENSIS apt AND apt
TITLE OF INVENTION: PROTEASE GENES AND ALKALINE PROTEASE DEFICIENT AND NEUTRAL
TITLE OF INVENTION: PROTEASE DEFICIENT B.T. STRAINS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Farnitch Schwabe Jacobs & Nadel, P.C.
STREET: 1601 Market Street, 36th floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19103-2398
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible


```

? Patent No. 5976545
? GENERAL INFORMATION:
? APPLICANT: Harris Ph.D., Jeffrey D.
? APPLICANT: Hsu, Kuang T.
? APPLICANT: Podolski, Joseph S.
? TITLE OF INVENTION: Pharmaceutical Compositions for
? TITLE OF INVENTION: Immunocontraception
? NUMBER OF SEQUENCES: 61
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
? ADDRESS: Borun
? STREET: 6300 Sears Tower, 233 South Wacker Drive
? CITY: Chicago
? STATE: Illinois
? COUNTRY: United States of America
? ZIP: 60606-6402
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? FILING DATE: 07-JUNE-95
? CLASSIFICATION: 514
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/149,223
? FILING DATE: 09-NOV-93
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/012,940
? FILING DATE: 29-JAN-93
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/973,341
? FILING DATE: 09-NOV-92
? ATTORNEY/AGENT INFORMATION:
? NAME: Clough, David W.
? REGISTRATION NUMBER: 36,107
? REFERENCE/DOCKET NUMBER: 32744
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 312/474-6653
? TELEFAX: 312/474-0448
? TELEX: 25-3856
? INFORMATION FOR SEQ ID NO: 16:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 570 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-484-158B-16

Query Match 0.5% Score 57; DB 2; Length 570;
Best Local Similarity 27.3%; Pred. No. 3,26+02;
Matches 12; Conservative 11; Mismatches 21; Indels 0; Gaps 0;

Q7 702 VTFHVLVSVSCSSSTKTHPEPAIVFSLIQKIKKLESVIT 745
:| | | | | | | | | | | | | | | | | | | | |
Db 308 ITRSIIFLRVSCSYSVSNAPFISVQVFTIPPPHUKTGHGPLT 351

RESULT 119
US-08-480-158A-16
? Sequence 16, Application US/08480150A
? Patent No. 5989550
? GENERAL INFORMATION:
? APPLICANT: Harris Ph.D., Jeffrey D.
? APPLICANT: Hsu, Kuang T.
? APPLICANT: Podolski, Joseph S.
? TITLE OF INVENTION: Materials and Methods for Immunocontraception
? NUMBER OF SEQUENCES: 59
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
? STREET: 6300 Sears Tower, 233 South Wacker Drive
? CITY: Chicago
? STATE: Illinois
? COUNTRY: United States of America
? ZIP: 60606-6402
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/480,150A
? FILING DATE: 07-JUN-1995
? CLASSIFICATION: 424
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/149,223

```



```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451,747
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/145,995
; FILING DATE: 29-OCT-1993
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAMS, GREGORY D.
; REGISTRATION NUMBER: 30901
; REFERENCE/DOCKET NUMBER: NEB-046-DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 927-5054
; TELEFAX: (508) 927-1705
; TELEX:
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 591 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-451-747-21

```

```

Query Match 0.5%; Score 57; DB 2; Length 591;
Best Local Similarity 29.9%; Pred. No. 3,440;
Matches 14; Conservative 10; Mismatches 23; Indels 0; Gaps 0;

```

```

QY 1348 KTKVMVTPALIQSDSGSIEVSRNVEIIVKIIISVFVDALPHVPEHR 1394
bb 208 KTKMKKKKKFSEVEVEFGIVVPTAFGLSSVKARDHDPDPHQ 254

```

```

RESULT 124
US-09-134-852-21
; Sequence 21, Application US/09134852
; Patent No. 6127148
; GENERAL INFORMATION:
; APPLICANT: CARLOW, CLOTILDE K.S.
; TITLE OF INVENTION: METHOD FOR IDENTIFYING ANTI-PARASITIC
; COMPOUNDS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/134,852
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/145,995
; FILING DATE: 29-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: PESNICK, DAVID S.
; REGISTRATION NUMBER: 34235
; REFERENCE/DOCKET NUMBER: 4406
; TELECOMMUNICATION INFORMATION:

```

```

RESULT 122
US-08-145-995A-21
; Sequence 21, Application US/08145995A
; Patent No. 5482850
; GENERAL INFORMATION:
; APPLICANT: CARLOW, CLOTILDE K.S.
; TITLE OF INVENTION: METHOD FOR IDENTIFYING ANTI-PARASITIC
; COMPOUNDS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/145,995A
; FILING DATE: 29-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PESNICK, DAVID S.
; REGISTRATION NUMBER: 34235
; REFERENCE/DOCKET NUMBER: 4406
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 523-4400
; TELEFAX: (617) 523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 591 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-145-995A-21

```

```

Query Match 0.5%; Score 57; DB 1; Length 591;
Best Local Similarity 29.8%; Pred. No. 3,440;
Matches 14; Conservative 10; Mismatches 23; Indels 0; Gaps 0;

```

```

QY 1348 KTKVMVTPALIQSDSGSIEVSRNVEIIVKIIISVFVDALPHVPEHR 1394
bb 208 KTKMKKKKKFSEVEVEFGIVVPTAFGLSSVKARDHDPDPHQ 254

```

```

RESULT 123
US-08-451-747-21
; Sequence 21, Application US/08451747
; Patent No. 5481107
; GENERAL INFORMATION:
; APPLICANT: CARLOW, CLOTILDE K.S.
; TITLE OF INVENTION: METHOD FOR IDENTIFYING ANTI-PARASITIC
; COMPOUNDS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GREGORY D. WILLIAMS; NEW ENGLAND BIOLOGICALS, INC.
; STREET: 32 TOZER ROAD
; CITY: BEVERLY
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 01915
; COMPUTER READABLE FORM:

```



```

1 CITY: Portland
2 STATE: Oregon
3 COUNTRY: U.S.A.
4 ZIP: 97204
5
6 COMPUTER READABLE FORM:
7 MEDIUM TYPE: Floppy disk
8 COMPUTER: IBM PC compatible
9 OPERATING SYSTEM: PC-DOS/MS-DOS
10 SOFTWARE: Patent in Release #1.0, Version #1.25
11
12 CURRENT APPLICATION DATA:
13 APPLICATION NUMBER: US/88/352 902D
14 FILING DATE: 09-Dec-1994
15 CLASSIFICATION: <unknown>
16 ATTORNEY/AGENT INFORMATION:
17 NAME: Van Kyselbergh, Pierre C.
18 REGISTRATION NUMBER: 33 957
19 REFERENCE/POCKET NUMBER: OHSI 406H
20 TELECOMMUNICATION INFORMATION:
21 TELEPHONE: (503) 224-6655
22 TELEFAX: (503) 295-6679
23 TELEX: 360619
24
25 INFORMATION FOR SEQ ID NO: 123:
26 SEQUENCE CHARACTERISTICS:
27 LENGTH: 770 amino acids
28 TYPE: amino acid
29 STRANDEDNESS: single
30 TOPOLOGY: linear
31 MOLECULE TYPE: protein
32
33 SEQUENCE DESCRIPTION: SEQ ID NO: 123:
34
35 US-08-253-002p-123
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
98
```

APPLICANT: Childval, Namit
TITLE OF INVENTION: Human E3 Ubiquitin Protein
Ligase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: ZENECA Pharmaceuticals, Inc.
STREET: 1800 Concord Pike
CITY: Wilmington
STATE: DE
COUNTRY: USA
ZIP: 19850-5437
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/070,060
FILING DATE: 30-APR-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60,073,839
FILING DATE: 05-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Higgins, Patrick H
REGISTRATION NUMBER: 39,704


```

US-09-603-665-5
Sequence 56, Application US-09-603-665-5
Patent No. 6,036,271
GENERAL INFORMATION:
APPLICANT: Tokai, Paul J.
APPLICANT: Mark, Michael R.
APPLICANT: Scudiero, David L.
APPLICANT: Baker, Kevin P.
APPLICANT: Barton, William F.
TITLE OF INVENTION: Protein, Tyrosine Kinases
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ATTORNEY/AGENT INFORMATION:
STREET: 400 Point St. Suite 1000
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 800 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.40
CURRENT APPLICATION DATA:
FILING DATE: 22 MAY 1996
CLASSIFICATION: C12N 1/21
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/170,558
FILING DATE: 20 DEC 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/152,563
FILING DATE: 25 NOV 1994
ATTORNEY/AGENT INFORMATION:
NAME: Basak, Janet E.
REGISTRATION NUMBER: 28,661
REFERENCE/KEY NUMBER: 85473
TELEPHONE: 415/225-1856
TELEFAX: 415/225-1856
TEXT: 91/071 7169
INF. PUBL. IN F. SEQ. ID NO.: 45
SEQUENCE CHARACTERISTICS:
LENGTH: 998 amino acids
TYPE: amino acid
TOPOLGY: linear
US-09-603-665-5
Query Match: 0.5% Score 57; DB 3; Length 886;
Best Local Similarity: 25.4%; Pred. No. 6,100,021;
Matches: 14; Conservative 17; Mismatches 29; Gaps 0;
29 123 HELDCHREERETGKLVYVHLLNLSKTRDHPDPLDMMHLEGLTLLDITGKGLD 1269
14 752 HYLRQGNRLQVLEFLDLSLSMSFVWNPDKLSEATLWTFNFKALDAPDQ 911
RESULT 147
US-09-603-665-5
Sequence 4, Application 09-603-665-5
Patent No. 6,060,449
GENERAL INFORMATION:
APPLICANT: Fisher, Robert
APPLICANT: Fisher, Yoram
APPLICANT: Zvi, Nami
APPLICANT: ZAMIR, David
TITLE OF INVENTION: A GENE FAMILY FROM THE 12 PUSARIAN RESISTANCE
TITLE OF INVENTION: LINES OF TOMATO AND USE THEREOF FOR TRANSFORMATION AND
TITLE OF INVENTION: SELECTIVE BREEDING OF TOMATO AND RELATED PLANTS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:

```

```

ADDRESS: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.40
CURRENT APPLICATION DATA:
ATTORNEY/AGENT INFORMATION:
FILING DATE: 09-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/06272
FILING DATE: 15-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 114,373
FILING DATE: 13-APR-1995
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1240 amino acids
TYPE: amino acid
TOPOLGY: linear
MOLECULE TYPE: protein
US-08-930-996A-4
Query Match: 0.5% Score 57; DB 3; Length 1240;
Best Local Similarity: 40.0%; Pred. No. 1,140,043;
Matches: 12; Conservative 5; Mismatches 13; Gaps 0;
97 2133 LPESHFLAEMETFEEDVDEDCGCTTGGC 2132
106 1034 LPELPSKELQGLNCFETEGELP-NLOKL 1063
RESULT 148
US-08-970-269A-29
Sequence 29, Application 08-08971269A
Patent No. 5976803
GENERAL INFORMATION:
APPLICANT: Kathryn Meek
TITLE OF INVENTION: Genetic Test For Equine Severe
TITLE OF INVENTION: Combined Immunodeficiency Disease
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESS: Dr. Benjamin A. Adler
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word for Macintosh
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-08-970,269A
FILING DATE: NO. 5976803 October 14, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Adler Ph.D., Benjamin A.
REGISTRATION NUMBER: 45,423
REFERENCE/KEY NUMBER: D8860
TELEPHONE: 713-777-2221
TELEFAX: 713-777-6908
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 2987 amino acids

```


MOLECULE TYPE: protein
HYPOTHETICAL: no
ANTI-SENSE: no
US-09-407-562-30

Query Match 0.5%; Score 57; DB 4; Length 3959;
Best Local Similarity 31.4%; Pred. No. 6.6e+03;
Matches 16; Conservative 29; Mismatches 29; Indels 0; Gaps 0;

QY 484 TSLTSLRLHLLASLFFETSYVQSGEMTMSKRVSLQINMPTPTPLTLLS 434
DB 131 TRTPFAASALSTSTPLVSLNVAELATGKKRKEGFGVGVCTLRNCS 181

RESULT 144
US-09-036-987A-5
Sequence 5, Application US/0903698/A
Patent No. 614526
GENERAL INFORMATION:
APPLICANT: Baltz, Richard H.
APPLICANT: Broughton, Mary C.
APPLICANT: Crawford, Kathryn P.
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Merlo, Donald J.
APPLICANT: Treadway, Patti J.
APPLICANT: Turner, Jan R.
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
FILING DATE: 1999-08-09
CURRENT FILING DATE: 1999-08-09
EARLIER APPLICATION NUMBER: US 09/36987
EARLIER FILING DATE: 1998-03-09
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 4928
TYPE: PRT
ORGANISM: Saccharopolyspora spinosa
US-09-370-700-5

Query Match 0.5%; Score 57; DB 4; Length 4928;
Best Local Similarity 27.8%; Pred. No. 8.6e+03;
Matches 10; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

QY 633 LHPLRGWEALENVKSTPKGKLGIVANOKMIELL 668
DB 4887 LHALLAKWDGARDGTARATSPQSLTAATDDEIFDLI 4922

RESULT 145
US-09-100-804-20
Sequence 20, Application US/09100804
Patent No. 6066472
GENERAL INFORMATION:
APPLICANT: GONZ, LEONEL JORGE
APPLICANT: SARAS, JAN
APPLICANT: CLAPSON-WELSH, LENA
APPLICANT: HELDIN, CARL-HENRIK
TITLE OF INVENTION: PRIMARY STRUCTURE AND FUNCTIONAL EXPRESSION OF NUCLEOSIDE SEQUENCES FOR NOVEL PROTEIN
TITLE OF INVENTION: TYROSINE PHOSPHATASES
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
STREET: 600 ATLANTIC AVENUE
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/09/100.804
APPLICATION DATE: 09-AUG-1996
FILING DATE: 09-AUG-1996
PRIOR APPLICATION NUMBER: US 08/115,573
FILING DATE: 01-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09943
FILING DATE: 01-SEP-1994
ATTORNEY/AGENT INFORMATION:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/596,291
FILING DATE: 09-AUG-1996
APPLICATION NUMBER: US 08/115,573
FILING DATE: 01-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09943
FILING DATE: 01-SEP-1994
ATTORNEY/AGENT INFORMATION:

US-09-370-700-5
Sequence 5, Application US/09370700
Patent No. 6274350
GENERAL INFORMATION:
APPLICANT: Baltz, Richard H.
APPLICANT: Broughton, Mary C.
APPLICANT: Crawford, Kathryn P.
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Treadway, Patti J.
APPLICANT: Turner, Jan R.
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
FILING DATE: 1999-08-09
CURRENT FILING DATE: 1999-08-09
EARLIER APPLICATION NUMBER: US 09/36987
EARLIER FILING DATE: 1998-03-09
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 4928
TYPE: PRT
ORGANISM: Saccharopolyspora spinosa
US-09-370-700-5

Query Match 0.5%; Score 57; DB 4; Length 4928;
Best Local Similarity 27.8%; Pred. No. 8.6e+03;
Matches 10; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

QY 633 LHPLRGWEALENVKSTPKGKLGIVANOKMIELL 668
DB 4887 LHALLAKWDGARDGTARATSPQSLTAATDDEIFDLI 4922

RESULT 146
US-09-100-804-20
Sequence 20, Application US/09100804
Patent No. 6066472
GENERAL INFORMATION:
APPLICANT: GONZ, LEONEL JORGE
APPLICANT: SARAS, JAN
APPLICANT: CLAPSON-WELSH, LENA
APPLICANT: HELDIN, CARL-HENRIK
TITLE OF INVENTION: PRIMARY STRUCTURE AND FUNCTIONAL EXPRESSION OF NUCLEOSIDE SEQUENCES FOR NOVEL PROTEIN
TITLE OF INVENTION: TYROSINE PHOSPHATASES
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
STREET: 600 ATLANTIC AVENUE
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/09/100.804
APPLICATION DATE: 09-AUG-1996
FILING DATE: 09-AUG-1996
PRIOR APPLICATION NUMBER: US 08/596,291
FILING DATE: 09-AUG-1996
APPLICATION NUMBER: US 08/115,573
FILING DATE: 01-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09943
FILING DATE: 01-SEP-1994
ATTORNEY/AGENT INFORMATION:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/596,291
FILING DATE: 09-AUG-1996
APPLICATION NUMBER: US 08/115,573
FILING DATE: 01-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09943
FILING DATE: 01-SEP-1994
ATTORNEY/AGENT INFORMATION:

MOLECULE TYPE: protein
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Rattus rattus
IMMEDIATE SOURCE:
CLONE: RHO20P
US-08-430-286A-11

Query Match 0.5%; Score 56; DB 4; Length 348;
Best Local Similarity 55.0%; Pred. No. 2,1e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
QY 1405 TLGAKEFLWILLILFEQYV 1424
111 1 11 111 111
Db 118 TLGEALMSLVVLAIRYV 147

RESULT 154
US-09-141-047-9
Sequence 9; Application US/09/141,047A
Patent No. 6043085
GENERAL INFORMATION:
APPLICANT: Walker, David H.
APPLICANT: Yu, Xue-Jie
TITLE OF INVENTION: Ehrlichia canis 120-kDa Immunodominant Antigenic
TITLE OF INVENTION: Protein Gene
FILE REFERENCE: D6143
CURRENT APPLICATION NUMBER: US/09/141,047A
CURRENT FILING DATE: 1998-08-27
NUMBER OF SEQ ID NOS: 15
SEQ ID NO 9
LENGTH: 406
TYPE: PNT
ORGANISM: Ehrlichia chaffeensis
FEATURE:
OTHER INFORMATION: Amino acid sequence of 120 kDa protein used to
OTHER INFORMATION: determine homology with E. canis 120 kDa protein.
PUBLICATION INFORMATION:
AUTHORS: Walker, David H.
AUTHORS: Yu, Xue-Jie
TITLE: Immunodominant 120 kDa Surface Exposed Adhesion
TITLE: Protein Genes of Ehrlichia chaffeensis
PATENT DOCUMENT NUMBER: US 08/656,034
PATENT FILING DATE: 1996-05-31
US-09-141-047-9

Query Match 0.5%; Score 56; DB 3; Length 406;
Best Local Similarity 37.5%; Pred. No. 2,7e+02;
Matches 15; Conservative 7; Mismatches 18; Indels 0; Gaps 0;
QY 2596 KUNIVILPISPELALMEDECEVPEQKTTQQLTV 2135
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 290 KEDELVQFSEPTVAESVSEVELETHLVLFN 26V 329

RESULT 155
US-08-385-186-2
Sequence 2; Application US/08385186
Patent No. 5744594
GENERAL INFORMATION:
APPLICANT: Adelman, John P
APPLICANT: Ashford, Michael J
APPLICANT: Bond, Chris T
TITLE OF INVENTION: DNA ENCODING ATP-SENSITIVE POTASSIUM
TITLE OF INVENTION: CHANNEL PROTEINS AND USES THEREOF
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Koorio and Crew
STREET: Stewart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California

COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/385,186
FILING DATE: 07-FEB-1995
CLASSIFICATION: 435
PFIOP APPLICATION DATA:
APPLICATION NUMBER: US 08/385,186
FILING DATE: 10-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/193,372
FILING DATE: 08-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Kruse, No. 5744594man J.
REGISTRATION NUMBER: 35,235
REFERENCE/DOCKET NUMBER: 14310 2 2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-385-186-2

Query Match 0.5%; Score 56; DB 1; Length 419;
Best Local Similarity 21.5%; Pred. No. 2,8e+02;
Matches 17; Conservative 19; Mismatches 43; Indels 0; Gaps 0;
QY 813 WTEFCEPEPEPVYHILITPEPM:NCADAVHEPV:KLETKVHEPEVPEVPEVSVIWT 872
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 20 NDKPKDKQAPVVPVPIATDPTPIAFGCKKPPQPMYKSGKCNVHHCNVQTYRYLSULET 79

QY 873 YCSLSNPLNCSVKTVLOT 891
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 80 TLVDLKWRFNLLVFTMYVT 98

RESULT 156
US-07-803-636A-2
Sequence 2; Application US/07803636A
Patent No. 6423428
GENERAL INFORMATION:
APPLICANT: MCGUIRE, TRAVIS C., TERRY F. McELWAIN, LANCE E. PERRYMAN,
APPLICANT: WILLIAM C. DAVIS
TITLE OF INVENTION: IMMUNIZATION AGAINST BABESIOSIS USING
TITLE OF INVENTION: IDENTIFIED SURFACE ANTIGENS OF BABESIA BIGEMINA AND SIMILAR
TITLE OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVID R. SALIWANCHIK
STREET: 2421 NW 41ST STREET, SUITE A-1
CITY: GAINESVILLE
STATE: FLORIDA
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/803,636A
FILING DATE: 19911206
CLASSIFICATION: 424


```

; STATE: WASHINGTON
; COUNTRY: USA
; ZIP: 98101-2347
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/537,540
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Shelton, Dennis K.
; REGISTRATION NUMBER: 26,997
; REFERENCE/DOCKET NUMBER: WSOPL11254
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206 695 1718
; TELEFAX: 206 224 0779
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 591 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US 08-537-540-4

Query Match 0.5%; Score 56; DB 2; Length 591;
Best Local Similarity 34.6%; Pred. No. 4.7e+02;
Matches 9; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 732 LQKKKKLESVITAVEIPSEWLEL 757
DB 224 LVKQINLSSIEFALFTHWVQM 249

RESULT 160
US-08-656-034-2
; Sequence 2, Application US/08656034
; Patent No. 6015691
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: Yu, Xue Jie
; TITLE OF INVENTION: IMMUNODOMINANT 120 KDA SURFACE-EXPOSED
; TITLE OF INVENTION: ADHESION PROTEIN GENES OF EHRlichia
; TITLE OF INVENTION: CHAFFENSIS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/656,034
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Cordier, Timothy S.
; REGISTRATION NUMBER: 38,414
; REFERENCE/DOCKET NUMBER: C138151
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 2:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 628 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-656-034-2

Query Match 0.5%; Score 56; DB 3; Length 628;
Best Local Similarity 37.5%; Pred. No. 5.2e+02;
Matches 15; Conservative 7; Mismatches 18; Indels 0; Gaps 0;

QY 2096 KHW:VLLFEETFLAPFMENGVVHPCQKTKIQIPV 2135
DB 210 KRFVTSQSSEPFVAESEVSKVQEETNEVLIKLDQV 249

RESULT 161
US-08-190-802A-63
; Sequence 63, Application US/08190802A
; Patent No. 5519003
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Ron, Dorit
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dellinger & Associates
; STREET: P.O. Box 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-0850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,802A
; FILING DATE: 01-FEB-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 8600-0139
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 713 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: T0P1, Fig. 46
; US-08-190-802A-63

Query Match 0.5%; Score 56; DB 1; Length 713;
Best Local Similarity 32.5%; Pred. No. 6.3e+02;
Matches 13; Conservative 11; Mismatches 16; Indels 0; Gaps 0;

QY 1159 EGVPIFLFPPPPAKPLCTVQKRPKPKKSKSDIESVQF 1198
DB 78 KHLKLGEGQFDHQIASITVQKQKQKQKQKQKQKQKQK 117

RESULT 162

```


100

STREET: ONE ARBOL PARK ROAD
 CITY: ARBOL PARK
 STATE: IL
 COUNTRY: USA
 ZIP: 60943-609
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 FILING DATE:
 APPLICATION NUMBER: P/US94/07480-1
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: PERMSKI, PRISCILLA F.
 REGISTRATION NUMBER: 33,209
 TELEPHONE: 708-947-6405
 TELEFAX: 708-947-6405
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3011 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULAR TYPE: protein
 POT-0594-07480-1

Query Match
 Best Local Similarity 34.68%, Score 56, ID 5, Length 3011
 Matches 9, Conservative 0, Mismatches 9, Indels 0, Gaps 0

27 1821 VLLPAIKTKYKQIENNNHMHMPEMS 1846
 13 1111 1111 1111

18 1745 VLPVAVQINWCKLEAFWBMWNPIS 1770

RESULT 180

US-09-603-665-5

Sequence 1, Application P/US94/07480-1

Patent No. 5428145

GENERAL INFORMATION:

APPLICANT: WATANABE, SHIN-ICHI

APPLICANT: YAMASHIRO, TOSIE

APPLICANT: USAMI, SUSUMU

APPLICANT: LEVARE, SUSUO

TITLE OF INVENTION: MAMMALIAN ENVELOPE GENES

TITLE OF INVENTION: ENVELOPE GENES

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: ARBOL PARK ROAD

STREET: ONE ARBOL PARK ROAD

CITY: ARBOL PARK

STATE: IL

COUNTRY: USA

ZIP: 60943-609

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

FILING DATE: 1992/08/07

APPLICATION NUMBER: P/US94/07480-1

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: PERMSKI, PRISCILLA F.

REGISTRATION NUMBER: 33,209

TELEPHONE: 708-947-6405

TELEFAX: 708-947-6405

TELEFAX: 708-938-2623
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3011 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULAR TYPE: protein
 POT-0595-01087-1

Query Match
 Best Local Similarity 34.68%, Score 56, ID 5, Length 3011
 Matches 9, Conservative 8, Mismatches 9, Indels 0, Gaps 0

27 1821 VLLPAIKTKYKQIENNNHMHMPEMS 1846
 13 1111 1111 1111

18 1745 VLPVAVQINWCKLEAFWBMWNPIS 1770

RESULT 180

US-07-925-695-5

Sequence 5, Application US/07925695

Patent No. 5428145

GENERAL INFORMATION:

APPLICANT: OKAMOTO, Hiroaki

APPLICANT: NAKAMURA, Tetsuo

TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOME

TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, ANTIGEN, ANTIBODY AND

TITLE OF INVENTION: DETECTION SYSTEMS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Reveridge, DeGrandi, Weillacher & Young

STREET: 1850 M Street, N.W., Suite 809

CITY: Washington

STATE: D.C.

COUNTRY: US

ZIP: 20006

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07925695

FILING DATE: 1992/08/07

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 28312791

FILING DATE: 09-AUG-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 660441/91

FILING DATE: 03-DEC-1991

ATTORNEY/AGENT INFORMATION:

NAME: Weillacher, Robert G.

REGISTRATION NUMBER: 20,541

TELEPHONE: (202) 659-2811

TELEFAX: (202) 659-1462

TELEX: W01 64470

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 3033 amino acids

TYPE: AMINO ACID

STRANDEDNESS: unknown

TOPOLOGY: linear

POT-07-925-695-5

Query Match
 Best Local Similarity 34.68%, Score 56, ID 5, Length 3033
 Matches 9, Conservative 9, Mismatches 9, Indels 0, Gaps 0


```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: 09/09177449
: FILING DATE: 1998-08-27
: CLASSIFICATION: C12N
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 09/0743096
: FILING DATE: August 1994
: ATTORNEY/AGENT INFORMATION:
: NAME: LINDA AXIMOVICH FLOYD
: REGISTRATION NUMBER: 347692
: REFERENCE/INVENTION NUMBER: BB 1001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (302) 992-9929
: TELEFAX: (302) 842-7994
: REFLEX: 65420
: INFORMATION FOR SEQ ID NO: 105:
: SEQUENCE CHARACTERISTICS
: LENGTH: 107 amino acids
: TYPE: AMINO ACID
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: PT: US-09-603-665-5

```

```

Query Match 0.58; Score 55; DB 5; Length 107;
Best Local Similarity 41.0%; Ident. No. 48;
Matches 16; Conservative 19; Mismatches 48; Indels 0; Gaps 0;

```

```

QY 742 LKRRKLEKLVAVLQPSWBIHMLKGLVWLVWVYVLSLQVAVVDSVFLVF 791

```

```

DB 95 VMEKRRKLEKLVAVLQPSWBIHMLKGLVWLVWVYVLSLQVAVVDSVFLVF 94

```

```

QY 792 SPPKLVAVLQPSWBIHMLKGLVWLVWVYVLSLQVAVVDSVFLVF 94

```

```

DB 95 SPPKLVAVLQPSWBIHMLKGLVWLVWVYVLSLQVAVVDSVFLVF 94

```

```

RESULT 164

```

```

US-09-141-145-2

```

```

Sequence 2; Application US/09141145

```

```

Patent No. 5961729

```

```

GENERAL INFORMATION:

```

```

APPLICANT: CHEN, Joon Yoon

```

```

ATTORNEY/AGENT INFORMATION:

```

```

NAME: LEE, Yoon Ben

```

```

REGISTRATION NUMBER: 1942741

```

```

TITLE OF INVENTION: Acid Isolated from Archaeopsis thaliana

```

```

FILE REFERENCE: 1942741

```

```

CURRENT APPLICATION NUMBER: US-09-141-145

```

```

FILING DATE: 1998-08-27

```

```

SEQUENCE INFORMATION:

```

```

SOFTWARE: WordPerfect 6.1/W/Indos

```

```

SEQ ID NO: 2

```

```

LENGTH: 235

```

```

TYPE: PR1

```

```

MOLECULE TYPE: protein

```

```

US-09-141-145-2

```

```

Query Match 0.58; Score 55; DB 2; Length 235;

```

```

Best Local Similarity 41.0%; Ident. No. 16;

```

```

Matches 11; Conservative 11; Mismatches 15; Indels 0; Gaps 0;

```

```

QY 997 GHLISFAPFISTAAVYVLTATLFFERKRRKRSQC 994

```

```

DB 997 GHLISFAPFISTAAVYVLTATLFFERKRRKRSQC 994

```

```

QY 997 GHLISFAPFISTAAVYVLTATLFFERKRRKRSQC 994

```

```

DB 997 GHLISFAPFISTAAVYVLTATLFFERKRRKRSQC 994

```

```

QY 997 GHLISFAPFISTAAVYVLTATLFFERKRRKRSQC 994

```

```

DB 997 GHLISFAPFISTAAVYVLTATLFFERKRRKRSQC 994

```

```

QY 997 GHLISFAPFISTAAVYVLTATLFFERKRRKRSQC 994

```

```

DB 997 GHLISFAPFISTAAVYVLTATLFFERKRRKRSQC 994

```

```

: APPLICANT: Colicelli, John D.
: TITLE OF INVENTION: CLONING BY COMPLEMENTATION AND RELATED
: TITLE OF INVENTION: PROCESSES
: NUMBER OF SEQUENCES: 88
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Loole, Gerstein, Murray & Boran
: STREET: 234 South Wacker Drive/5400 Sears Tower
: CITY: Chicago
: STATE: Illinois
: COUNTRY: United States of America
: ZIP: 60606 6402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/474,379C
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/511,715
: FILING DATE: 20-APR-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/206,188
: FILING DATE: 01-MAR-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/588,352
: FILING DATE: 19-APR-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Cloghly, David W.
: REGISTRATION NUMBER: 36,107
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (312) 474-6300
: TELEFAX: (312) 474-0448
: INFORMATION FOR SEQ ID NO: 88:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 346 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-474-379C-88

```

```

Query Match 0.58; Score 55; DB 2; Length 346;
Best Local Similarity 40.0%; Ident. No. 2,80-02;
Matches 12; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

```

```

QY 2022 PVEGLENRGGKFKFQVTKHLIPCTAQ 2051

```

```

DB 152 PLPDKTQFFHLEEEEEELSMAGHICTAQ 181

```

```

RESULT 186

```

```

US-09-177-349-3

```

```

Sequence 3; Application US/09177449

```

```

Patent No. 6258201

```

```

GENERAL INFORMATION:

```

```

APPLICANT: Alland, David

```

```

ATTORNEY/AGENT INFORMATION:

```

```

NAME: Jacobs Jr., William X.

```

```

TITLE OF INVENTION: Inhibitors of Mycobacteria and Methods

```

```

FILE REFERENCE: 96700/491

```

```

CURRENT APPLICATION NUMBER: US/09/177,349

```

```

FILING DATE: 1998-10-24

```

```

NUMBER OF SEQ ID NOS: 14

```

```

SOFTWARE: Patent In Ver. 2.0

```

```

SEQ ID NO: 3

```

```

LENGTH: 479

```

```

TYPE: PR1

```

```

MOLECULE TYPE: protein

```

```

US-08-474-379C-88

```

```

Query Match 0.58; Score 55; DB 2; Length 479;

```

```

Best Local Similarity 40.0%; Ident. No. 2,80-02;

```

```

Matches 12; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

```

```

QY 2022 PVEGLENRGGKFKFQVTKHLIPCTAQ 2051

```

```

DB 152 PLPDKTQFFHLEEEEEELSMAGHICTAQ 181

```

```

: APPLICANT: Colicelli, John D.
: TITLE OF INVENTION: CLONING BY COMPLEMENTATION AND RELATED
: TITLE OF INVENTION: PROCESSES
: NUMBER OF SEQUENCES: 88
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Loole, Gerstein, Murray & Boran
: STREET: 234 South Wacker Drive/5400 Sears Tower
: CITY: Chicago
: STATE: Illinois
: COUNTRY: United States of America
: ZIP: 60606 6402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/474,379C
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/511,715
: FILING DATE: 20-APR-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/206,188
: FILING DATE: 01-MAR-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/588,352
: FILING DATE: 19-APR-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Cloghly, David W.
: REGISTRATION NUMBER: 36,107
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (312) 474-6300
: TELEFAX: (312) 474-0448
: INFORMATION FOR SEQ ID NO: 88:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 346 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-474-379C-88

```

```

Query Match 0.58; Score 55; DB 2; Length 479;

```

```

Best Local Similarity 40.0%; Ident. No. 2,80-02;

```

```

Matches 12; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

```

```

QY 2022 PVEGLENRGGKFKFQVTKHLIPCTAQ 2051

```

```

DB 152 PLPDKTQFFHLEEEEEELSMAGHICTAQ 181

```

```

QY 2022 PVEGLENRGGKFKFQVTKHLIPCTAQ 2051

```

```

DB 152 PLPDKTQFFHLEEEEEELSMAGHICTAQ 181

```

```

QY 2022 PVEGLENRGGKFKFQVTKHLIPCTAQ 2051

```

```

DB 152 PLPDKTQFFHLEEEEEELSMAGHICTAQ 181

```

```

QY 2022 PVEGLENRGGKFKFQVTKHLIPCTAQ 2051

```

```

DB 152 PLPDKTQFFHLEEEEEELSMAGHICTAQ 181

```

```

QY 2022 PVEGLENRGGKFKFQVTKHLIPCTAQ 2051

```

```

DB 152 PLPDKTQFFHLEEEEEELSMAGHICTAQ 181

```

```

US-08-206-188B-12
? Sequence 12, Application US/08206188B
? Patent No. 6100025
? GENERAL INFORMATION:
? APPLICANT: Widler, Michael H.
? ADDRESSEE: Colicelli, John J.
? TITLE OF INVENTION: Cloning by Complementation and Related
? NUMBER OF SEQUENCES: 84
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
? STREET: 6300 Sears Tower, 233 South Wacker Drive
? CITY: Chicago
? STATE: Illinois
? COUNTRY: United States of America
? ZIP: 60606-6402
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1 0, Version #1 35
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/206,188B
? FILING DATE: 01-MAR-1994
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/511,715
? FILING DATE: 20-APR-1990
? ATTORNEY/AGENT INFORMATION:
? NAME: Crough, David W.
? REGISTRATION NUMBER: 36137
? TELEPHONE: 312-474-6400
? TELEFAX: 312-474-0448
? TELEX: 25-3856
? INFORMATION FOR SEQ ID NO: 12:
? LENGTH: 898 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
US-08-206-188B-12

```

```

Query Match 0.5%; Score 55; DB 3; Length 898;
Best local similarity 40.0%; Pred. No. 1.2e+03;
Matches 12; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

```

```

QY 2022 PLVQVLEPFGGFEKQFVYKRIQCTAQ 2051
Db 676 PLPKQFQELTLEEEEREISMAQIPICTAQ 705

```

```

RESULT 194
US-07-688-352B-12
? Sequence 12, Application US/07688352B
? Patent No. 5527896
? GENERAL INFORMATION:
? APPLICANT: Widler, Michael H.
? ADDRESSEE: Colicelli, John J.
? TITLE OF INVENTION: Cloning by Complementation and Related
? NUMBER OF SEQUENCES: 57
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
? ADDRESSEE: Ricknell
? STREET: Two First National Plaza, 20 South Clark
? CITY: Chicago
? STATE: Illinois
? COUNTRY: USA
? ZIP: 60603
? COMPUTER READABLE FORM:

```

```

? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1 0, Version #1 25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/07/688,352C
? FILING DATE: 19910419
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/511,715
? FILING DATE: 20-APR-1990
? ATTORNEY/AGENT INFORMATION:
? NAME: Borun, Michael P.
? REGISTRATION NUMBER: 25447
? REFERENCE/DOCKET NUMBER: 27805/30197
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (312) 346-5750
? TELEFAX: (312) 984-9740
? TELEX: 25-3856
? INFORMATION FOR SEQ ID NO: 12:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 900 amino acids
? TYPE: AMINO ACID
? TOPOLOGY: linear
? MOLECULE TYPE: protein
US-07-688-352B-12

```

```

Query Match 0.5%; Score 55; DB 1; Length 900;
Best local similarity 40.0%; Pred. No. 1.3e+03;
Matches 12; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

```

```

QY 2022 PLVQVLEPFGGFEKQFVYKRIQCTAQ 2051
Db 678 PLPKQFQELTLEEEEREISMAQIPICTAQ 707

```

```

RESULT 199
US-08-968-752B-6
? Sequence 6, Application US/08968752B
? Patent No. 6043073
? GENERAL INFORMATION:
? APPLICANT: Frohman, Michael A.
? APPLICANT: Morris, Andrew
? TITLE OF INVENTION: DNA Sequences
? NUMBER OF SEQUENCES: 8
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: GNYX Pharmaceuticals, Inc.
? STREET: 3031 Research Drive
? CITY: Richmond
? STATE: California
? COUNTRY: USA
? ZIP: 94806
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1 0, Version #1 30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/968,752B
? FILING DATE: 13-AUG-1997
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/025,469
? FILING DATE: 05-SEP-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Giotta, Gregory J.
? REGISTRATION NUMBER: 32,028
? REFERENCE/DOCKET NUMBER: GNYX2004
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 510-222-9700
? TELEFAX: 510-222-9758

```

1 INFORMATION FOR SEQ ID NO: 61:
 2 SEQUENCE CHARACTERISTICS:
 3 LENGTH: 552 amino acids
 4 TYPE: amino acid
 5 TOPology: linear
 6 MOLECULE TYPE: protein
 7 US 09-603-665-5

Query Match: 0.54; Score 55; DB 4; Length 942;
 Best Local Similarity: 26.00; Pred. No. 1.4e+04;
 Matches: 14; Conservation: 8; Mismatches: 29; Indels: 0; Gaps: 0;
 27 14 15 VENVVHLSKGRERFLSVGVMSGLSSNNEKRVVVSQDEFLKCLEE 1519
 28 62 11NAVYHLIRSQHLYTENFFISISGRIVKNGVGVIVRIVKRAHEQ 671

RESULT 260
 DB of 560,005
 Sequence 7, Application: US/08560005
 Patent No. 6,916,54
 GENERAL INFORMATION:
 APPLICANT: 302, David A.
 APPLICANT: Williams, Lewis I.
 APPLICANT: Jefferson, Anne Bennett
 APPLICANT: Marotus, Philip W.
 TITLE OF INVENTION: No. 60016441 GRD2 Association Protein and Nucleic
 TITLE OF INVENTION: Acidus Encoding Thereof
 NUMBER OF SEQUENCES: 10
 REFERENCE ADDRESS:
 ADDRESS: Townsend and Townsend and Crew
 STREET: One Market Plaza, Steuart Tower, Suite 2000
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94105

COMPUTER PROGRAM: FIRM
 MEDIUM TYPE: Floppy Disk
 MEDIUM: FIRM PC Format File
 OPERATING SYSTEM: PC/OS/MS DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/560,005
 FILING DATE:
 CLASSIFICATION: 4.05

ALL KEY/AGENT INFORMATION:
 NAME: 302, Karen W.
 PETS: 302, Karen W.
 REFERENCE/KEY NUMBER: 29,781
 REFERENCE/KEY NUMBER: 2,678,024,000
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415, 426-2400
 TELEFAX: 415, 426-2422
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 568 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FEATURES:

NAME/KEY: Protein
 LOCATION: 1-568
 OTHER INFORMATION: Note: "act1"

US 09-603-665-5

Query Match: 0.54; Score 55; DB 4; Length 968;
 Best Local Similarity: 41.84; Pred. No. 1.4e+04;
 Matches: 14; Conservation: 19; Mismatches: 20; Indels: 0; Gaps: 0;
 27 47 10TDRRLDLDSEDFEAFIFSLAKTLERSVQIKAVNKQL 90
 28 48 10TDRRLDLDSEDFEAFIFSLAKTLERSVQIKAVNKQL 90

Search completed: November 6, 2001, 12:20:56
 Job time: 253 sec

TeXmacs version 4.5
Copyright (c) 1998 - 2002 Compton Ltd.

Mathematical Population Genetics

Run on: November 6, 2001, 12:20:43; Search time 56.04 Seconds
(without alignments)
5153.742 Million cell updates

1st Lt. US-04-603-665-5
 10807
 1st Lt. US-04-603-665-5
 10807
 1st Lt. US-04-603-665-5
 10807

Simulation table: `HL.SIM.2`
`data = list(), vapour = 0`

Searched: 425,46 scops, 1240627 positions

Total number of fits satisfying chosen parameters: 42505

The figure illustrates the mechanism of action of efavirenz against HIV-1. It depicts the process of reverse transcription where HIV-1 RNA is converted into DNA by the enzyme Reverse Transcriptase (RT). Efavirenz is shown as a red molecule that binds to RT, inhibiting its activity. This inhibition blocks the synthesis of the viral DNA intermediate. Consequently, the viral RNA is degraded, preventing the integration of viral DNA into the host genome. Without integrated DNA, no new HIV-1 particles can be produced.

[illegible]

THE HISTORY OF THE

- 1: *Sp. arcticeps* *
- 2: *Sp. bartolici* *
- 3: *Sp. berti* *
- 4: *Sp. berti* *
- 5: *Sp. inversipalpis* *
- 6: *Sp. mammali* *
- 7: *Sp. miki* *
- 8: *Sp. orfium* (c)
- 9: *Sp. phaeo* *
- 10: *Sp. plant* *
- 11: *Sp. ruber* *
- 12: *Sp. unclavati* *
- 13: *Sp. zettleri* *
- 14: *Sp. zettleri* *

Fred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Position	* QuinT		Longitude	B	U	Description
	Score	Match				
1	4201	58.9	8.47	6	Q9M44	Q9M44 maruca fasc
2	4063	28.3	4.11	4	Q9H23	Q9H23 homo sapien
3	3776	16.5	3.39	4	Q9N23	Q9N23 homo sapien
4	511	4.7	2.06	5	Q9VW75	Q9VW75 drosophila
5	199	1.8	1.49	3	Q6U179	Q6U179 schizosacch
6	14	1.4	1.55	5	Q24495	Q24495 caenorhabd
7	7	0.7	1.54	3	Q9P1M8	Q9P1M8 schizosacch
8	76	0.7	1.46	3	Q9Y634	Q9Y634 schizosacch
9	75	0.7	2.45	10	Q9SE64	Q9SE64 eucumis sat
10	75	0.7	9.60	5	Q9V900	Q9V900 drosophila
11	74	0.7	1.66	5	Q9NM7	Q9NM7 leishmania
12	74	2.7	4.08	10	Q9S030	Q9S030 arabidopsis
13	73	0.7	4.48	13	Q73665	Q73665 galus gall
14	73	0.7	4.70	4	Q9H9N5	Q9H9N5 homo sapien
15	73	0.7	4.79	4	Q9H439	Q9H439 homo sapien
16	73	0.7	4.90	2	Q93778	Q93778 trepanema p
17	73	0.7	11.27	14	Q9YV16	Q9YV16 melanoplas
18	72	0.7	4.55	13	Q73587	Q73587 scyllorhina
19	72	0.7	5.63	19	Q6U700	Q6U700 arabidopsis

Davidson K., Doup I.E., Dowdes M., Dugan Rocha S., Eankov P., Eam
 Durbin K.J., Evangelista C.C., Ferray C., Ferreira S., Fleischmann W.,

Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
-1- SIMILARITY: TO YEAST YJL109C.

```

EMBL: A02628.1; AAC16972.1;
InfoProc: TRF000497;
DB: PROSITE; position: real_posval: 1
KW: Hypothetical protein; Transmembrane;
FT: TRANSMEM 254-274 potential;
FT: DOMAIN 1728-1821 P-GLY ASP;
SQ: SEQUENCE 1649 AA; 166409 MW; 6041D47A12521649 CRC64;

Query Match:
Best Local Similarity: 45.9%; Score 199; DB 4; Length 1649;
Matches: 4; Conservative: 21; Mismatches: 57; Indels: 0; Gaps: 0;

Q7 Q7 SLIDKREAAATHHETAFASDTHEELHLDPSPGFEFAHLESQAETLERSVQITAVN 87
||||| ||| | | | | | | | | | | | | | | | |
DB SLIDKRVAAAMERLEIIV AVSDFHLLAYHRLKLATLTLGGQGVQTVRLRTER 87
||| | | | | | | | | | | | | | | | |
Q7 KK KATENISSEFLHSVFFLHWAGTLEWLPEDTHLYNQUSLAVLVYHETIEVEV 147
||| | | | | | | | | | | | | | | | |
DB EKIPDQVGLHETLAVERTEFNALKEVLEWLPERSTREYVDEEILSLPDRWFARI 147
||| | | | | | | | | | | | | | | | |

Q7 148 I 148
DB 148 I 148

RESULT *
QZ495 PRELIMINARY: PRI: 1650 AA.
QZ495
DI 01-NOV-1996 (TRF000497) Created)
DI 01-MAY-1996 (TRF000497) (1, Last sequence update)
DI 01-MAY-2000 (TRF000497) (1, Last annotation update)
CDDED FOR BY C. ELEGANS (CNA YK483.5).
ZK483.1
Caenorhabditis elegans.
Bakteriyota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
Rhabditidae; Polioderinae; Caenorhabditis.
NBI_LINID 6239;
[1]
SEQUENCE FROM N.A.
MEDLINE 9415070; PubMed 7906498;
Wilson R., Ainsworth R., Anderson K., Baynes C., Bicks M.,
Reinhold J., Burton J., Connell M., Copsey L., Cooper J., Coulson A.,
Craxton M., Dear S., Du Z., Durbin R., Fazelto A., Fulton L.,
Garner A., Green P., Hastings T., Hillier L., Jir M., Johnston L.,
Jones M., Kerstan J., Kingston J., Laister N., Latreille P.,
Lachlan J., Lloyd C., Nemuray A., Mortimore B., O'Callaghan M.,
Parsons J., Percy C., Ripken J., Rooper A., Saunders D., Showkoon P.,
Smallman N., Smith A., Sommer E., Staden R., Sulston J.,
Thierry-Mieg F., Thomas R., Vardiu M., Varian H., Waterston P.,
Watson A., Weinstock J., Wilkinson-Spoat J., Wodhnan P.,
# 2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans?
Nature 408:42-46(1994).
[2]
SEQUENCE FROM N.A.
Johnsen P.;
Submitted (Jan 1997) to the EMBL/GenBank/DDBJ databases.
[3]
SEQUENCE FROM N.A.
Waterston P.;
Submitted (Oct 1995) to the EMBL/GenBank/DDBJ databases
EMBL: U02833; AAA05811; ;
InfoProc: TRF000497;
DB: PROSITE; position: real_posval: 1
FT: TRANSMEM 17000524;
FT: STIE; PS00043; HH.GTK.FAMILY; UNKN-WN 1;
DB PROSITE; PS00640; HH.GTK.FAMILY; UNKN-WN 1;
SQ: SEQUENCE 1530 AA; 185426 MW; A143D38739FD29 CR764;

Query Match:
Best Local Similarity: 1.5%; Score 163; DB 5; Length 1650;
Matches: 1; Conservative: 0; Mismatches: 15; Indels: 0; Gaps: 0;

```

	Matches	31: Conservative	24: Mismatches	43: Indels	0: Gaps
QY	64	QFAPLEPSAKRTIPSSVQFAVAKFQIDENISFTTHSPVELFKALCKGLFWLTHREHI	124		
DB	65	EPANDESEKVEFVSRMLPKANERKQTEKLLHLSHYLQHWQGVFLPLHYQI	124		
QY	124	HLYNQSLACVLPYHTRIFVVIQLL	151		
DB	125	YSPNAETILLTFLPFHETKVSRLRL	152		
RESULT	7				
Q9P7M8					
ID	Q9P7M8	PRELIMINARY:	PR1: 1564 AA.		
AC	Q9P7M8				
DT	01-DEC-2000 (TREMBLrel. 15, Created)				
DT	01-DEC-2000 (TREMBLrel. 15, Last sequence update)				
DT	01-DEC-2000 (TREMBLrel. 15, Last annotation update)				
DE	NUCLEOPORIN.				
GN	NUPI84.				
OS	Schizosaccharomyces pombe (fission yeast).				
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;				
OC	Schizosaccharomycetales; Schizosaccharomycetaceae;				
OC	Schizosaccharomycetes.				
OX	NCBI_TaxID=4896;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=972H;				
RA	Seeger K., Harris D., Lync M., Rajandream M.A., Barrell B.G.;				
DR	Submitted (FEB 2000) to the EMBL/GenBank/CCRF databases.				
DR	EMBL: AL157917; CAB76031.1; -				
KW	Porin.				
SC	SEQUENCE 1564 AA; 176462 MW; 248PFAFF39C40E7 CRC64;				
Query Match	0.7%;	Score 76;	EB 4;	Length 1564;	
Best local similarity	41.9%;	Prod. No. 13;			
Matches	18; Conservative	6; Mismatches	19; Indels	0; Gaps	0;
QY	1709	NVIGSALICIAEVTISIEALATPQPSLMPSTLITMKNSHIV	1751		
DB	886	NCLPRLCTAPIQLISALLAPWQSEFSTIAYMINSTDIV	928		
RESULT	8				
Q9Y8G4					
ID	Q9Y8G4	PRELIMINARY:	PR1: 1628 AA.		
AC	Q9Y8G4				
DT	01-NOV-1999 (TREMBLrel. 12, Created)				
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)				
DT	01-MAY-2000 (TREMBLrel. 13, Last annotation update)				
DE	NUCLEOPORIN.				
GN	NUPI84.				
OS	Schizosaccharomyces pombe (fission yeast).				
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;				
OC	Schizosaccharomycetales; Schizosaccharomycetaceae;				
OC	Schizosaccharomycetes.				
OX	NCBI_TaxID=4896;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=972;				
RA	MEDLINE=99318821, PubMed=10488805;				
RA	Whalen W.A., Yoon J.H., Shen R., Dhar K.;				
FT	"Regulation of mRNA export by nutritional status in fission yeast.";				
FL	Genetics 152:827-838(1999).				
DR	EMBL: AF055035; AAD43830.1; -				
KW	Porin.				
SC	SEQUENCE 1628 AA; 184236 MW; AEC72819FFFF40DB7 CRC64;				
Query Match	0.7%;	Score 76;	DB 4;	Length 1628;	
Best local similarity	41.9%;	Prod. No. 14;			
Matches	18; Conservative	6; Mismatches	19; Indels	0; Gaps	0;


```

Query Match 0.7%; Score 73; DB 4; length 470;
Best Local Similarity 30.7%; Pred. No. 8.2;
Matches 23, Conservative 12; Mismatches 40; Indels 0; Gaps 0;

QY 2065 LNYGULLETRHSSPKVEEAMITVLALAEFKENYIVGLFESTFTFLAFMDPRTVEVHQ 2124
DB 194 LQELGOLKEAENLAAPRAVDVAATLARLIDELSKLSNEETAFILKKVIEELRELQAO 253
QY 2125 CQKTCQGLFTVIGEP 2139
DB 254 LQPGGQGVEMDMGSKP 268

RESULT 14
QY 083778 PRELIMINARY: PRT: 490 AA.
AC 083778
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE SERINE-TYPE D-ALA-D-ALA CARBOXYPEPTIDASE (DACC).
TP0800.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NICHOLS;
RX MEDLINE=98332770; PubMed=9665870;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
Dodson R., Cowlin M., Hickey E.K., Clayton P., Ketchum K.A.,
Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
Khailak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
McDonald L., Artach P., Bowman K., Cotton M.D., Fujii C., Garland S.,
Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
Venter J.C.;
RA "Complete genome sequence of Treponema pallidum, the syphilis
spirochete.";
RL Science 281:375-388(1998).
EMBL: AE001251; AAC65769.1; -
DR MEROPS: S11.001; -
DR TIGR: TP0800; -
DR InterPro: IP001967; -
DR Pfam: PF00768; peptidase S11; 1.
DR PRINIS: PR00725; DALACPTASE1.
DR CARBOXYPEPTIDASE.
KW SEQUENCE 490 AA; 52809 MW; 7289A98ACEFACE3 CRC64;

Query Match 0.7%; Score 73; DB 2; length 490;

QY 291 TKFSLIKDGLSCLIVLGRQKSPESIGKKPFPHLCNVPDLITILHGISYDVSPLLRYM 350
DB 370 TDVETASDALPCAIVLGSKRPGALRPILHPSTCTSCPVLTNPGRITRISITFALPPLLRAP 429
QY 351 LPHLVVSIHHVTGE 365
DB 430 LQETDVGIFAHVLDE 444

RESULT 17
QY 09YVT6 PRELIMINARY: PRT: 1127 AA.
AC 09YVT6
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE ORF MSV156 HYPOTHETICAL PROPTIN.
GN MSV156.
OS Melanoplus sanguinipes entomopoxvirus (MSEPV).
OC Entomopoxvirus B.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Entomopoxvirinae;
OX NCBI_TaxID=83191;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TUCSON;
RX MEDLINE=99102412; PubMed=9847359;
RA Afonso C.L., Tulman E.P., Lu Z., Oma F., Kutish G.F., Rock D.L.;
RT "The genome of Melanoplus sanguinipes entomopoxvirus.";
RL J. Virol. 73:533-552(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TUCSON;
RA Afonso C.L., Tulman E.P., Lu Z., Oma F., Kutish G.F., Rock D.L.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF064860; AAC97677.1; -
SQ SEQUENCE 1127 AA; 134265 MW; F185D4LD5A3FE7D1 CRC64;

Query Match 0.7%; Score 73; DB 14; length 1127;
Best Local Similarity 21.3%; Pred. No. 22;
Matches 19, Conservative 22; Mismatches 48; Indels 0; Gaps 0;

QY 1461 GSENNILQYLLKPEEKDEIDKAVSPNKSEQEMLOVNVETHTSKLPHFKFLSVSF 1520
DB 164 ELLLELLKLVSSLSKSLGIEIYKHHHREIEPKRGRVKKLIRPEL;DFLIFPS 223
QY 1521 MSOLLSSNNFLKVVESGGPPIKGLPER 1549
DB 124 EKLIPEQEELNFTIFKEQENIKFKINOK 250

RESULT 18
QY 073587 PRELIMINARY: PRT: 455 AA.
AC 073587
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE DESMIN.
OS Scyllorhinus stellaris.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Carchariniiformes;
OC Scyllorhinidae; Scyllorhinus.
OX NCBI_TaxID=68454;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MUSCLE;
RA Schultess J., Loebbecke A., Schaffeld M., Lieb B., Markl J.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO THE INTERMEDIATE FILAMENT FAMILY.

```

DB	EMBL	Y15064	AA: 547	1	Score	72	DB	14	Length	455
DB	InterPro	IPRO01664	1							
DB	Flam	PF00048	Filament	1						
DB	PF03112	PS00226	IP	1						
DB	KW	Coiled coil	Heptad repeat pattern: intermediate filament							
DB	SEQUENCE	455 AA: 52230 MW: 83352	AFCE8546E9	CR664						
Query Match										
Best Local Similarity	40.7%									
Matches	23	Conservation	11	Mismatches	41	Identities	0	Gaps		
DB	EMBL	U01111	AA: 111	1	Score	72	DB	14	Length	214
DB	InterPro	IPRO01664	1							
DB	Flam	PF00048	Filament	1						
DB	PF03112	PS00226	IP	1						
DB	KW	Coiled coil	Heptad repeat pattern: intermediate filament							
DB	SEQUENCE	455 AA: 52230 MW: 83352	AFCE8546E9	CR664						
Query Match										
Best Local Similarity	40.7%									
Matches	23	Conservation	11	Mismatches	41	Identities	0	Gaps		
DB	EMBL	U01111	AA: 111	1	Score	72	DB	14	Length	214
DB	InterPro	IPRO01664	1							
DB	Flam	PF00048	Filament	1						
DB	PF03112	PS00226	IP	1						
DB	KW	Coiled coil	Heptad repeat pattern: intermediate filament							
DB	SEQUENCE	455 AA: 52230 MW: 83352	AFCE8546E9	CR664						
Query Match										
Best Local Similarity	40.7%									
Matches	23	Conservation	11	Mismatches	41	Identities	0	Gaps		
DB	EMBL	U01111	AA: 111	1	Score	72	DB	14	Length	214
DB	InterPro	IPRO01664	1							
DB	Flam	PF00048	Filament	1						
DB	PF03112	PS00226	IP	1						
DB	KW	Coiled coil	Heptad repeat pattern: intermediate filament							
DB	SEQUENCE	455 AA: 52230 MW: 83352	AFCE8546E9	CR664						
Query Match										
Best Local Similarity	40.7%									
Matches	23	Conservation	11	Mismatches	41	Identities	0	Gaps		
DB	EMBL	U01111	AA: 111	1	Score	72	DB	14	Length	214
DB	InterPro	IPRO01664	1							
DB	Flam	PF00048	Filament	1						
DB	PF03112	PS00226	IP	1						
DB	KW	Coiled coil	Heptad repeat pattern: intermediate filament							
DB	SEQUENCE	455 AA: 52230 MW: 83352	AFCE8546E9	CR664						
Query Match										
Best Local Similarity	40.7%									
Matches	23	Conservation	11	Mismatches	41	Identities	0	Gaps		
DB	EMBL	U01111	AA: 111	1	Score	72	DB	14	Length	214
DB	InterPro	IPRO01664	1							
DB	Flam	PF00048	Filament	1						
DB	PF03112	PS00226	IP	1						
DB	KW	Coiled coil	Heptad repeat pattern: intermediate filament							
DB	SEQUENCE	455 AA: 52230 MW: 83352	AFCE8546E9	CR664						
Query Match										
Best Local Similarity	40.7%									
Matches	23	Conservation	11	Mismatches	41	Identities	0	Gaps		
DB	EMBL	U01111	AA: 111	1	Score	72	DB	14	Length	214
DB	InterPro	IPRO01664	1							
DB	Flam	PF00048	Filament	1						
DB	PF03112	PS00226	IP	1						

[illegible]

Query Match 0.7%; Score 71; DB 10; Length 1269;
Best Local Similarity 37.0%; Prod. No. 46;
Matches 17; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN-LANUSBERG ERECTA; TISSUE=FLOWER;
RC Collier M.E., Saunders M.J.;
RC Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF004556; AAD01202.1;
DR EMBL: AF004556; AAD01202.1;
SQ SEQUENCE 275 AA; 32273 MW; ED962625E552A5 FPC64;

Query Match 0.6%; Score 69; DB 10; Length 275;
Best Local Similarity 23.9%; Pred. No. 15;
Matches 17; Conservative 21; Mismatches 33; Indels 0; Gaps 0;

QY 1217 KVFIPSPQIIIVPTLWNI SPCIFPI PQGFENMEYTKQI IISGLINICQKISDYGCKIPKD 1276
DB 43 EETLEKFTVTETKMPVLNINERILEAGCSMESSIISLINAMFHLERQVSVSDGIECPG 102

QY 1277 ILDERKPNVEL 1287
DB 103 VLTVDKIKFEL 113

RESULT 29
ID 083089 PRELIMINARY; PRT: 202 AA.
AC 083089;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE CONSERVED HYPOTHETICAL PROTEIN.
GN TP00050.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=NICHOLS;
RC Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RC Dodson R., Gwinn M., Hickley E.K., Clayton R., Ketchum K.A.,
RC Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RC Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RC McDonald L., Artach P., Rowman C., Cotton M.D., Fujii C., Garland S.,
RC Hatch K., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RC Venter J.C.;
RC "Complete genome sequence of Treponema pallidum, the syphilis
RC spirochete."
RC Science 281:375-388(1998).
RL EMBL: AE001190; AA065046.1;
DR TIGR: TP00050;
DR InterPro: IPR000836;
DR Pfam: PF00156; Pribosyltran; 1.
DR Transferrase.
KW SEQUENCE 202 AA; 23208 MW; F703AF702ARQQLVA CRC64;
SQ

Query Match 0.6%; Score 69; DB 2; Length 262;
Best Local Similarity 27.4%; Pred. No. 14;
Matches 26; Conservative 12; Mismatches 41; Indels 0; Gaps 0;

QY 719 KIESVITAVELPSEWHIEMIDRGIPVELWAHYHVNSTORVAVESVFLVFLSKFIY 798
DB 55 KAPGPIIYAAVVAHSHTLPKFRVIVSVGWTYPTFVAVGKVLIVGLFDSGATHYIA 114

QY 799 ALKAKKSPKGTI 811
DB 115 SLLMKGLARGDI 127

Query Match 0.6%; Score 68; DB 1; Length 286;
Best Local Similarity 27.3%; Pred. No. 21;
Matches 21; Conservative 9; Mismatches 47; Indels 0; Gaps 0;

QY 2049 IAGE-SVAMADUSLWKPI NYOII LKTRDSSEKVRFAALITVLALAKLKERYIVLLPSIP 2108
DB 33 LADLDVEAPNDHLLGLVGLQNFRRVHQFMKFNVAKCICKRKAFCVCEHAIIVTLKGTIP 92

QY 2109 LAFLEMEDECEVEHQHQC 2125
DB 93 FLMPITLCSGCRACRIVC 109

RESULT 31
ID Q9V0C8 PRELIMINARY; PRT: 409 AA.
AC Q9V0C8;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE ORF MSV194 ALI MOTIF GENE FAMILY PROTEIN.
GN MSV194.
OS Melanoplus sanguinipes entomopoxvirus (MSEPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Entomopoxvirinae;
OX Entomopoxvirus B.
OX NCBI_TaxID=83191;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=TUCSON;
RC MEDLINE=99102512; PubMed=9847359;
RC AF063866; Tulman E.P., Tu Z., Oms W., Forish G.F., Rock D.L.;
RC "The genome of Melanoplus sanguinipes entomopoxvirus."
RC J. Virol. 73:533-552(1999).
RN [2];
RP SEQUENCE FROM N.A.
RC STRAIN=TUCSON;
RC Afonso C.I., Tulman E.P., Tu Z., Oms W., Forish G.F., Rock D.L.;
RC Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF063866; AAC97765.1;
DR InterPro: IPR003497;
DR Pfam: PF02498; BRO; 1.
DR SEQUENCE 409 AA; 48676 MW; A062DEF999B47D57 CRC64;
SQ

Query Match 0.6%; Score 68; DB 1; Length 286;
Best Local Similarity 27.3%; Pred. No. 21;
Matches 21; Conservative 9; Mismatches 47; Indels 0; Gaps 0;

QY 2049 IAGE-SVAMADUSLWKPI NYOII LKTRDSSEKVRFAALITVLALAKLKERYIVLLPSIP 2108
DB 33 LADLDVEAPNDHLLGLVGLQNFRRVHQFMKFNVAKCICKRKAFCVCEHAIIVTLKGTIP 92

QY 2109 LAFLEMEDECEVEHQHQC 2125
DB 93 FLMPITLCSGCRACRIVC 109

RESULT 31
ID Q9V0C8 PRELIMINARY; PRT: 409 AA.
AC Q9V0C8;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE ORF MSV194 ALI MOTIF GENE FAMILY PROTEIN.
GN MSV194.
OS Melanoplus sanguinipes entomopoxvirus (MSEPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Entomopoxvirinae;
OX Entomopoxvirus B.
OX NCBI_TaxID=83191;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=TUCSON;
RC MEDLINE=99102512; PubMed=9847359;
RC AF063866; Tulman E.P., Tu Z., Oms W., Forish G.F., Rock D.L.;
RC "The genome of Melanoplus sanguinipes entomopoxvirus."
RC J. Virol. 73:533-552(1999).
RN [2];
RP SEQUENCE FROM N.A.
RC STRAIN=TUCSON;
RC Afonso C.I., Tulman E.P., Tu Z., Oms W., Forish G.F., Rock D.L.;
RC Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF063866; AAC97765.1;
DR InterPro: IPR003497;
DR Pfam: PF02498; BRO; 1.
DR SEQUENCE 409 AA; 48676 MW; A062DEF999B47D57 CRC64;
SQ

[illegible]

DR HSPD: P02933; IRXD;
DR InterPro: IPR000410; -;
DR InterPro: IPR000598; -;
DR InterPro: IPR004594; -;
DR Pfam: PF00512; signal; 1;
DR Pfam: PF00672; DUF5; 1;
DR SMART: SM00387; HAIPase_C1;
KW Kinase; Phosphorylation; Serine transferase.
SQ SEQUENCE 454 AA; 50650 MW; 9f5cfad6b2b3961 CRC64;

Query Match
Best Local Similarity 21.3%; Pred. No. 45;
Matches 13; Conservative 22; Mismatches 29; Indels 0; Gaps 0;

CY 172 PR ISYVDELISQVHLEKRISEWASGZAMFPLSLSEELALILAPVLPAFLPEE LGL
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 245 PEVLFEETAEISEAHHMFPTNQIALAFSGHGLEVELEKFTITLKAAPAYMTLSVVQ 424
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 1832 Q 1832
DB 325 R 325

RESULT 34
QNTX6
ID QNTX6 PRELIMINARY; PRI: 405 AA;
AC QNTX6;
PI 01-MAY-2000 (TRENDEL) 12; Last sequence update;
DI 01-MAY-2000 (TRENDEL) 13; Last sequence update;
DT 01-JUN-2000 (TRENDEL) 14; Last annotation update;
DE HYPOTHEtical 52.0 KDA PROTEIN.
GN HRP.
OS Cyanidium caldarium (Galdieria sulphuraria).
OC Chloroplast.
EC Eukaryota; Rhodophyta; Rhodospirillum rubrum; Porphyridiales; Porphyridiaceae;
NC Cyanidium.
CX NCBI_TaxID 2771;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RKI;
RA Gleeson G., Rosenthal A., Valentin K.;
ET "Reconstruction of the ancient red algal plastid genome: Structure,
FC gene repertoire, and synteny";
BU Submitted (Nov-1999) to the EMBL/GenBank/Genbank databases.
BK EMBO: A022195; AAI1292.1;
DD InterPro: IPR002917;
DF Pfam: PF01926; WWK_HSR1; 1;
DR Hypothetical protein; Chloroplast;
KW SEQUENCE 405 AA; 52904 MW; 81229519e1c61652 CRC64;
SQ SEQUENCE 405 AA; 52904 MW; 81229519e1c61652 CRC64;

Query Match
Best Local Similarity 34.2%; Pred. No. 46;
Matches 13; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

CY 408 SSQFFMSSKKVSLIFQFHLPLRLLESKPRTLLWVLE 445
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 315 SSEDVSIFNKMLMKKNKFVILLNLDKXVCRIEIE 402
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 34
Q90441
ID Q90441 PRELIMINARY; PRI: 473 AA;
AC Q90441;
DT 01-NOV-1996 (TRENDEL) 01; Created;
DI 01-MAY-2000 (TRENDEL) 13; Last sequence update;
PI 01-MAY-2000 (TRENDEL) 14; Last annotation update;
DE DESMIN
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
NC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysii;
CY Cypriniformes; Cyprinidae; Poeciliinae; Danio;
KW

1000


```

DR HSSP: P11746; 1MMN.
DR Mendel: 32687; Matdo:MADS; 42687.
DR InterPro: IPR002100; -.
DR InterPro: IPR002487; -.
DR Pfam: PF00319; SRP-TF; 1.
DR Pfam: PF01486; K-box; 1.
DR PRINTS: PR00404; MADS DOMAIN.
DR PROSITE: PS00350; MADS_BOX_1; 1.
DR PROSITE: PS00066; MADS_BOX_2; 1.
DR SMART: SM00432; MADS; 1.
DR DNA-binding: Nuclear protein, Transcription regulation.
DR KW DNA-binding: Nuclear protein; Transcription regulation.
DR SEQUENCE 246 AA; 28278 MW; 641F35E21C887FA1 CF64;

Query Match
Best local similarity 18.9%; Score 67; DB 10; Length 246;
Matches 18; Conservative 26; Mismatches 51; Indels 0; Gaps 0;

QY 1819 PVLLPAIKTKYQIEKNNKNNHMGPEMSIIORHIGXMKKFEITSHQSQTAFPLEALDPR 1878
DB 84 PAKELSSYREYMKLGKPYESLQPTNPILGDIQPIINKELEQLFQLEGSLKQVRSTK 143

QY 1879 AQHSENDERVGKTENCIIDCLVAMVVKLSEVIFR 1913
DB 144 TOYMLDQLSDLQNKQELLLEANKDLTKMLDEISSR 178

QY 1879 AQHSENDERVGKTENCIIDCLVAMVVKLSEVIFR 1913
DB 144 TOYMLDQLSDLQNKQELLLEANKDLTKMLDEISSR 178

RESULT 41
ID Q9SHA6 PRELIMINARY; PRT: 246 AA.
AC Q9SHA6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE MADS BOX PROTEIN.
DE MDADSB PROTEIN.
GN MDADSB.
OS Malus domestica (Apple) (Malus sylvestris).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Eudicotyledons; core eudicots; Rosidae; eucosids I;
OC Eucosids; Rosaceae; Malus.
OX NCBI_TaxID=3750;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-CV. GRANNY SMITH;
RA YAO J., Dong Y., Kaurtheden A., Morris B.,
RA "Seven apple MADS-box genes are expressed in different parts of
RA fruit.";
RA Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION FACTORS.
EMBL: AJ001681; CAA04319.1;
DR HSSP: P11746; 1MMN.
DR InterPro: IPR00730; -.
DR InterPro: IPR002100; -.
DR InterPro: IPR002487; -.
DR Pfam: PF00319; SRP-TF; 1.
DR Pfam: PF01486; K-box; 1.
DR PRINTS: PR00404; MADS DOMAIN.
DR PROSITE: PS002673; -.
DR PROSITE: PS00350; MADS_BOX_1; 1.
DR PROSITE: PS00066; MADS_BOX_2; 1.
DR SMART: SM00432; MADS; 1.
DR DNA-binding: Nuclear protein; Transcription regulation.
DR KW DNA-binding: Nuclear protein; Transcription regulation.
DR SEQUENCE 246 AA; 28250 MW; 761DF23E8B87RA2 CF64;

Query Match
Best local similarity 18.9%; Score 67; DB 10; Length 246;
Matches 18; Conservative 26; Mismatches 51; Indels 0; Gaps 0;

QY 1819 PVLLPAIKTKYQIEKNNKNNHMGPEMSIIORHIGXMKKFEITSHQSQTAFPLEALDPR 1878
DB 84 PAKELSSYREYMKLGKPYESLQPTNPILGDIQPIINKELEQLFQLEGSLKQVRSTK 143

```

```

QY 1879 AQHSENDERVGKTENCIIDCLVAMVVKLSEVIFR 1913
DB 144 TOYMLDQLSDLQNKQELLLEANKDLTKMLDEISSR 178

RESULT 42
ID Q9LFU6 PRELIMINARY; PRT: 262 AA.
AC Q9LFU6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE MADS BOX PROTEIN AGL2.
GN FL4PR 180
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Eudicotyledons; core eudicots; Rosidae; eucosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Sato S., Nakamura Y., Kaneke T., Kato T., Asamizu F., Kotani H.,
RA Tabata S., Mewes H.W., Ruid S., Lemcke K., Mayer K.F.X.;
RA Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Arabidopsis sequencing project;
RA Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION FACTORS.
EMBL: AL391144; CAC01779.1; -.
DR InterPro: IPR002100; -.
DR InterPro: IPR002487; -.
DR Pfam: PF00319; SRP-TF; 1.
DR Pfam: PF01486; K-box; 1.
DR PRINTS: PR00404; MADS DOMAIN.
DR PROSITE: PS00350; MADS_BOX_1; 1.
DR PROSITE: PS00066; MADS_BOX_2; 1.
DR SMART: SM00432; MADS; 1.
DR DNA-binding: Nuclear protein; Transcription regulation.
DR KW DNA-binding: Nuclear protein; Transcription regulation.
DR SEQUENCE 262 AA; 29914 MW; 6990357E8B87RA2 CF64;

Query Match
Best local similarity 18.1%; Score 67; DB 10; Length 262;
Matches 15; Conservative 25; Mismatches 43; Indels 0; Gaps 0;

QY 1828 KTYKQIEKNNKNNHMGPEMSIIORHIGXMKKFEITSHQSQTAFPLEALDPR 1887
DB 93 FEYLRKQPYENLQPCQPNLLQFNLQPLNSKFEIQLERQDLSIKQVRSTKTYMIDLS 152

QY 1888 EVGKTENCIDCLVAMVVKLSEV 1910
DB 153 DLONKEQMLLETNRALAMKIDDM 175

RESULT 43
ID Q9FLV7 PRELIMINARY; PRT: 349 AA.
AC Q9FLV7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE GENOMIC DNA, CHROMOSOME 5, PL CLONE:MZF18.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Eudicotyledons; core eudicots; Rosidae; eucosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-COLUMBIA;

```

```

PX MEDLINE 9427054; PubMed 9428582;
RA Saito S, Kuroda T, Kotani H, Nakamura Y, Asaike E, Miyajima N,
RA Tabata S;
PI "Genetical analysis of Arabidopsis thaliana chromosome 5, IV.
PI Sequence features of the regions of 1,456,315 bp covered by nineteen
PI physically assigned pl and IAC clones."
RA INA Res. 5:41-54(1998).
DR EMBL: AH070594; BAB04728.1;
SQ SEQUENCE 344 AA: 69783 MW: 34598695841029 CR654;

Query Match 0.6%; Score 67; DB 10; Length 349;
Best Local Similarity 33.9%; Pred. No. 45;
Matches 19; Conservative 9; Mismatches 28; Indels 0; Gaps 0;

CY 1476 EKEELIPRAVSNKSESDPEMLQVNVFTHSKQDHPHFVSVPMSQLSSNNF 1530
1111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
10 30 EEEEDRRFIVSEFFEFKFFDFFDFFKFFSSSVSEVSEVNIIVSLSSTLQNGF 31

RESULT 44
Q15867
ID Q15867 PRELIMINARY; PRI: 454 AA.
A: Q15867;
DI 01-NOV-1996 (TrEMBLrel. 01; Created)
DI 01-NOV-1996 (TrEMBLrel. 12; Last sequence update)
DI 01-MAR-2001 (TrEMBLrel. 16; Last annotation update)
DE VIMENLIN (SPERMATOPHYTES).
GN VIM.
OS Homo Sapiens (Homo).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN 111
SQ SEQUENCE FROM N.A.
PX MEDLINE 9422008; PubMed 9371605;
RA Perron J, Lillienbaum A, Vasseux M, Paulin B;
PI "Nucleotide sequence of the human vimentin gene and regulation of its
PI transcription in tissues and cultured cells."
RA Gene 627:16(1998).
CC 1 SIMILARITY: TO THE INTERMEDIATE FILAMENT FAMILY.
DR EMBL: M18895; AAAA1281.2; 1.
DR EMBL: M18888; AAAA1281.2; JOINED.
DR EMBL: M18889; AAAA1281.2; JOINED.
DR EMBL: M18890; AAAA1281.2; 1.
DR EMBL: M18891; AAAA1281.2; JOINED.
DR EMBL: M18892; AAAA1281.2; JOINED.
DR EMBL: M18893; AAAA1281.2; JOINED.
DR EMBL: M18894; AAAA1281.2; JOINED.
DR TrEMBL: P1800104; 1.
DR Pfam: PF00038; 1.
DR ProSITE: PS00226; 1.
KW Coiled coil; Heptad repeat pattern; Intermediate filament.
FI N-TER.
SQ SEQUENCE 344 AA: 41562 MW: 90764003099707 CR664;

Query Match 0.6%; Score 62; DB 4; Length 354;
Best Local Similarity 26.6%; Pred. No. 47;
Matches 21; Conservative 17; Mismatches 41; Indels 0; Gaps 0;

QY 2061 LKPKNYQLHLETSSTPVGAAATVVALAGALGGRATVLEFLPTALMMSERE 2120
111 111 111 111 111 111 111 111 111 111 111 111 111 111
10 7A LKPKGFEMEGQEFAPSTLSEGVWVNASLAPDPEPKVSEQLFAPKSGEEFGE 132

QY 2121 VEHQVKTUQLTGLDEP 2149
111 111 111 111 111 111 111 111 111 111 111 111 111 111
10 13A LKAGLGEHWGVNVVSKP 151

RESULT 44
ID 44 PRELIMINARY; PRI: 466 AA.

```

```

AC 045910;
DI 01-JUN-1998 (TrEMBLrel. 06; Created)
DI 01-JUN-1998 (TrEMBLrel. 06; Last sequence update)
DI 01-MAR-2001 (TrEMBLrel. 16; Last annotation update)
DE Y17D7B.1 PROTEIN.
GN Y17D7B.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN 111
SQ SEQUENCE FROM N.A.
RA Korshaw J;
PI Subunit 1 (XAN 1548) to the YML-5 subunit, TUBB1 databases.
RN 121
SQ SEQUENCE FROM N.A.
RX MEDLINE 94150718; PubMed 7906398;
RA Williams R, Ainsworth R, Andersen K, Payton C, Parks M,
RA Bonfield J, Burton J, Connell M, Copsey J, Cooper J, Coulson A,
RA Craxton M, Dear S, Du Z, Durbin R, Favetto A, Fulton L,
RA Gardner A, Green P, Hawkins T, Hillier L, Jier M, Johnston L,
RA Jones M, Kersey J, Kirsten J, Laister N, Latreille P,
RA Lightning J, Lloyd C, McMurray A, Mortimore B, O'Callaghan M,
RA Parsons J, Perry C, Riiken L, Rogers A, Saunders D, Showkeen R,
RA Smalton N, Smith A, Sonhammer E, Staden R, Sulston J,
RA Thierry-Mieg J, Thomas K, Vaudin M, Vaughan K, Waterston R,
RA Watson A, Wellstock T, Wilkins-Sprout J, Wohlhueter P;
PI "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
PI elegans."
RA Nature 368:32-38(1994).
CC 1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC 1- SIMILARITY: TO C4 TYPE STEROID RECEPTOR ZINC FINGER FAMILY.
DR EMBL: A2321472; CAA16299.1; 1.
DR HSSP: P10828; 2NLL.
DR InterPro: IPR000536; 1.
DR InterPro: IPR001628; 1.
DR Pfam: PF00105; ZF-C4; 1.
DR SMART: SM00430; HOL1; 1.
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
SQ SEQUENCE 486 AA: 56732 MW: 55546660820846 CR664;

Query Match 0.6%; Score 67; DB 5; Length 486;
Best Local Similarity 23.8%; Pred. No. 51;
Matches 10; Conservative 14; Mismatches 18; Indels 0; Gaps 0;

QY 933 FEMMLNCADAVHFFVMELEFEVHEVFEVFEVFEVFEVFEVFEVFEVFEV 874
111 111 111 111 111 111 111 111 111 111 111 111 111 111
10 276 FEFVFGGKMIVFVIFVFEVFEVFEVFEVFEVFEVFEVFEVFEVFEV 411

RESULT 46
Q9RFNO
ID Q9RFNO PRELIMINARY; PRI: 668 AA.
AC Q9RFNO;
DI 01-MAY-2000 (TrEMBLrel. 13; Created)
DI 01-MAY-2000 (TrEMBLrel. 13; Last sequence update)
DI 01-MAR-2001 (TrEMBLrel. 16; Last annotation update)
DE BETA-GALACTOSIDASE HGAR.
GN HGAR.
OS Carnobacterium piscicola.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
OC Carnobacterium.
OX NCBI_TaxID=2751;
RN 111
SQ SEQUENCE FROM N.A.
RC STRAIN-BA;
RX MEDLINE 20049983; PubMed 10584002;
RA Coombs J.M., Brechley J.E.;
PI "Biochemical and phylogenetic analysis of a cold-active beta-
PI galactosidase from the lactic acid bacterium Carnobacterium piscicola
PI strain BA."

```

Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ikegawa C., Jalali M., Kalish F., Karpen G.H., Ko Y., Kennison J.A., Ketchum K.A., Kimble B.E., Kivira C.B., Kraft C., Krasilev S., Kulp P., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattoti B., McIntosh T.C., McLeod M.P., McPherson D., Merkushin S., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.N., Moy M., Murphy R., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Sinden-Klamos L., Simpson M., Skupski M.P., Smith T., Spier E.C., Spradling A.C., Stapleton M., Strong R., Sun E., Swirsky S., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R. F., Zaveri J.S., Zhang M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu G., Zhu X., Smith H.O., Gibbs P.A., Myers E.W., Rubin G.M., Venter J.C.;
"the genome sequence of *Drosophila melanogaster*";
Science 287:2185-2195(2000).
EMBL: AE003638; AAF53237.1; -.
HSSP: P42345; IEPF.
FlyBase: FBgn032466; G55092.
InterPro: IPR000403; -.
InterPro: IPR001440; -.
InterPro: IPR003151; -.
InterPro: IPR003152; -.
Pfam: PF00454; P13_P14_kinase; 1.
Pfam: PF00515; TPR; 1.
Pfam: PF02259; FAT; 1.
Pfam: PF02660; FATC; 1.
PROSITE: PS00915; P13_4_KINASE_1; 1.
PROSITE: PS00916; P13_4_KINASE_2; 1.
PROSITE: PS00920; P13_4_KINASE_3; 1.
SMART: SM00146; P13Kc; 1.
SEQUENCE 2470 AA; 281031 MW; 507812RC07C7FF9 CRC64;

Query Match: C.6%; Score 65; DB 2; Length 433;
Best Local Similarity 27.9%; Pred. No. 82;
Matches 15; Conservative 15; Mismatches 26; Indels 0; Gaps
151; ACCESSION: S040607; FILE: PROKAP1.CT; SIZE: 1000; MD5SUM: E119B

1. *What is the purpose of this study?*

$$Z_1^2 = Z_1^2(\text{Fe}, \text{N}, \text{L}) = 2.4119, \quad 6\text{Å}; \quad Z_1^2(\text{Fe}, \text{O}, \text{L}) = 2.4104, \quad 6\text{Å};$$

Key words: The culture nucleotide sequences of three hepatitis B virus


```

SU SEQUENCE      405 AA;   34069 MW    770DE459D1EE5C09  47 364;

Query Match:
Best local Similarity 28.1%; Score 44; ID# 2; Length 405;
Matches 18; Conservative 12; Mismatches 34; Indels 0; Gaps 0;

QY 448 LPFLAIPPLPGLHGVSTISGKYLALASLSSLSLNLNLAAPVLLAMNLESLME 507
II : : : I : : : : : : : : : : : : : : : : : : : : : : : :
DB 17 LEALVRLNFLEVALFELPELVAVPEEVLGFEETLLEALLTAYSLTHLEEDENLR 76
II : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 508 ISKE 511
I :
DB 77 EIKD 80

RESULT 84
Q9R7H6
IT Q949HC PRELIMINARY; PFI; 463 AA.
AC Q9R7H6;
DT 01 MAY 2000 (TrEMBLrel_18, (rotated)
BT 01 MAY 2000 (TrEMBLrel_18, last sequence update)
DT 01 MAY 2000 (TrEMBLrel_18, last annotation update)
DE HYPOTHEICAL 41-9 KDA PROTEIN (FRAGMENT).
OS Borrelia burgdorferi ( Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrellia.
GN NCHI_taxid:139;
UN NCHI_taxid:139;
KN KN
RC SEQUENCE FROM N.A.
RF STRAIN N40;
RA Feng S., Chen G., Barthold S.;
RT "G46 clone from Borrelia burgdorferi.";
RL Submitted (MAY-1997) to the FMBI/GenBank/DBJ databases.
DR EMBL: AF005055; AAA01225.1; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 463 AA; 41909 MW; D8330154940942FA CR064;

```

Query Match	0.68	Score	64	DB	2	Length	36
Post Local	Similarity	34.1%	Prod.	No.	90		
Matches	14	conservative	9	MisMatches	18	Indels	0
						Gaps	0

QY	61	SFEQFFAPLFSJIAKTIKRSVQTCAVKNKKQLDENISLSLIHL	101
	11	: :	
DB	231	SPASYRINFNFKTYOAFERKEEELANKMMIMGIIYIYLIGF	271

RESULT	85	
Q9KSS2		
1D	Q9KSS2	PRELIMINARY; PRI: 416 AA.
AC	Q9KSS2;	
DT	01-OCT-2000 (TREMBLREL_15, Created)	
DI	01-OCT-2000 (TREMBLREL_15, Last sequence update)	
DT	01-MAR-2001 (TREMBLREL_16, Last annotation update)	
DE	NIP5 PELATEL PROTEIN	
GN	VC1184.	
OS	Vibrio cholerae.	
OC	Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibr.	
OX	NCBI_TaxID 666;	
EN	[1]	
RP	SEQUENCE FROM N.A.	
SC	STEATH P.; T.F. SUGA; /	SECURITY-E 01;
FX	MEDLINE=20406833; PubMed=10952401;	
EA	Heidelberg J.F., Eisen J.A., Nelson W.N., Clayton K.A., Gwinn M.L.,	
EA	Ledson P.J., Hatt H.H., Hickey E.K., Peterson J.D., Mayday L.A.,	
EA	Gill S.F., Nelson K.E., Read T.P., Tettelin H., Richardson D.,	
EA	Ercegovac M.D., Vamathevan J., Bass S., Hin H., Brar H., Sellers P.	
EA	McDonald L., Ufferkack T., Flisvianant R.D., Norman W.C., White O.,	
EA	Salzberg S.L., Smith H.O., Colwell R.K., Mekalanos J.J., Venter J.C.,	
EA	Fraser C.M.;	
RT	"DNA sequence of both chromosomes of the cholera pathogen Vibrio	
RT	cholerae."	

RA Bird C.?
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO THE INTERMEDIATE FILAMENT FAMILY.
DR EMBL: AL133415; CAB87963.1; -
DR InterPro: IPR001664; -
DR Pfam: PF00038; filament; 1.
DR PRINTS: PR01248; TYPEKERATIN.
DR PROSITE: PS00226; IF; 1.
KW Coiled coil; Heptad repeat pattern; Intermediate filament.
SQ SEQUENCE 466 AA; 55651 MW; BAB54026056015A CRC64;

Query Match 0.6%; Score 64; DB 4; Length 466;
Best Local Similarity 32.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 12; Mismatches 22; Indels 0; Gaps 0;
QY 2090 ALAKLKNYVILLPSIPHLAEDECEVEHQCQKTIQOLETVLGEP 2139
DB 214 SLARLDIERKVESLQERLAFIKKIHHEETQELQAOIQFHQVQIDVDVSKP 263

RESULT 88
Q9NUX6 PRELIMINARY; PRT; 553 AA.
AC Q9NUX6;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE CDNA FLJ11071 F15, CLONE P1ACF1604937, MODERATELY SIMILAR TO SPL-110 PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RX NCBI_TaxID=9606;
RN [1]
RS SEQUENCE FROM N.A.
RA TISSUE=PLACENTA;
RA Isogai I., Ota I., Hayashi K., Sugiyama T., Ohsaki T., Suzuki Y.,
RA Nishikawa I., Nagai K., Suquho S., Aotsuka S., Yashikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;
RP "NPOD human cDNA sequencing project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK001933; BAA91966.1; -
DR InterPro: IPR001680; -
DR Pfam: PF00400; WD40; 7.
DR PRINTS: PR00646; F-box; 1.
DR PROSITE: PR00320; GPROTEINRPT.
DR PROSITE: PS00181; FBOX; 1.
DR SMART: SM00676; WL_KPEALS, UNKNOWNL5.
KW Repeat; WD repeat.
SQ SEQUENCE 553 AA; 62280 MW; CAB90271096A3F3 CRC64;

Query Match 0.6%; Score 64; DB 4; Length 553;
Best Local Similarity 27.5%; Pred. No. 1.5e+02;
Matches 19; Conservative 10; Mismatches 40; Indels 0; Gaps 0;
QY 50 IGLERLGLGSLSFQFELVLFGLAFTLEFVYVYVAVYKQLEINISLFLHISPYEYLPK 109
DB 91 SGFKLLALDELIDSCPTQVRIMQVIEFTQFQRFISLLPEALYVLSFPPKDTICA 150
QY 110 AQKCLEWLI 118
DB 151 AOTCYWR 159
RESULT 89
Q9DGF4 PRELIMINARY; PRT; 740 AA.
ID Q9DGF4

RA Nature 406:477-493(2000).
RL EMBL: AF004198; AAF94343.1; -
CC TIGR: VC1184; -
DR InterPro: IPR000192; -
DR Pfam: PF00266; aminotran_5; 1.
DR SEQUENCE 416 AA; 45878 MW; 7511ABAL60C3208B CRC64;
Query Match 0.6%; Score 64; DB 2; Length 416;
Best Local Similarity 24.6%; Pred. No. 1.1e+02;
Matches 14; Conservative 13; Mismatches 30; Indels 0; Gaps 0;
QY 1337 LDQYFQVQVINKVAVIHALGQSGDSLESNVWEIVVKEIISVVDALPHVPEH 1393
DB 154 LDQAHLSLNPPTPIVAVTFASNTSIVETFAIETAAIQVAMVYVAVVILPHH 210

RESULT 86
Q58581 PRELIMINARY; PRT; 450 AA.
AC Q58581;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE 450AA LONG HYDROTHERMAL PMO PROTEIN.
DR PH0851.
OS Pyrococcus horikoshii
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
RX NCBI_TaxID=53953;
RN [1]
RS SEQUENCE FROM N.A.
RA STRAIN=014;
RA MEDLINE=96344137; PubMed=9676194;
RA Kawarabayashi Y., Swids M., Horikawa H., Nakawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka K., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudo Y., Yamazaki J., Kushiida N., Ouchi A.,
RA Aoki Y., Yoshikawa Y., Nakamura Y., Pebb F.T., Horikoshi K.,
RA Masurhi Y., Shiroya H., Kikuchi H.,
RA "Complete sequence and gene organization of the genome of a hyper-
KT thermophilic archaeobacterium, Pyrococcus horikoshii OT3."
RL DNA Res. 5:55-76(1998).
DR EMBL: AP000003; BAA29445.1; -
DR InterPro: IPR000051; -
DR Pfam: PF01189; Noll_Nop2_Sunp; 1.
DR SEQUENCE 450 AA; 51637 MW; D2952324B1EFAAF3C CRC64;

Query Match 0.6%; Score 64; DB 1; Length 450;
Best Local Similarity 27.0%; Pred. No. 1.1e+02;
Matches 13; Conservative 14; Mismatches 25; Indels 0; Gaps 0;
QY 348 PYMPEHVVSTTHVTCPTTMDQIVKRHEALITKLSKNNILHLSL 399
DB 154 KYLAUSWLIERKVGHDGDETEDFKSVNKRHWISIRVNTLKAVDEVEL 205
RESULT 87
Q9NTM3 PRELIMINARY; PRT; 466 AA.
AC Q9NTM3;
DT 01-OCT-2000 (TREMBl rel. 15, Created)
DT 01-OCT-2000 (TREMBl rel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBl rel. 16, Last annotation update)
DE BA124N14.1 (VIMENTIN).
GN VIM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RX NCBI_TaxID=9606;
RN [1]
RS SEQUENCE FROM N.A.

RESULT	%	
Q84893		
ID	Q84893	PRELIMINARY; PRT; 878 AA.
AC	Q84893	
DT	01-NOV-1996	(TREMBLrel. 01, Created)
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)
DT	01-OCT-2000	(TREMBLrel. 15, Last annotation update)
DE	CARSID PROTEINS (FRAGMENTS).	
OS	Human poliovirus 3.	
OC	Viruses; ssRNA positive-strand viruses, no DNA stage, Picornaviridae.	
OC	Enterovirus.	
OX	NCBI_Taxid:12086;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN SAUKETT A.	
RX	MEDLINE:94160574; PubMed 8116248;	
RT	HuCV31-94; Kinnunen, L., Poyry T., Laakkonen L., Koivainen M.,	

SEQUENCE FROM N.A.
 STRAIN=A3(2);
 MEDLINE=97000351; PubMed=8843436;
 PA Redenbach M., Krieser H.M., Benaparte D., Eichner J.,
 PA Kinashi H., Hopwood D.A.;
 KT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RT Mol. Microbiol. 21:77-96(1996).
 FL FMBL; A1160431; CAR77429.1; -;
 DR InterPro; IPK00722; -;
 DR Pfam; PF00623; RNA_pol_A; 1.
 KW DNA-directed RNA polymerase.
 FT NON_TER 1059 1059
 FT SEQUENCE 1059 AA; 118914 MW; 852436A859106E3C CRC64;

Query Match 0.68; Score 64; DB 5; Length 1154;
 Best Local Similarity 29.48; Ref. No. 340-02;
 Matches 20; Conservative 10; Mismatches 48; Indels 0; Gaps 0;
 QY 360 NNIDHILASIFFEYVTSYSSQPFEMISNKVSIINFGELDIPIIFSKYDPTIDVIFEFHLK 449
 DB 1081 NNAVHYMKILPARKTIFISQIMENISFVFIQIDPQKDFEYLKLIVE 1140
 QY 450 ELAULKQ 457
 DB 1141 YVADMFQ 1148
 RESULT 97
 QYVW49
 ID QYVW49 PRELIMINARY PRI: 1779 AA.
 AC QYVW49;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
 DE CG8334 PROTEIN (FRAGMENT).
 GN CG8334.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Erosophilidae; Drosophila.
 CX NCBI_TaxID 7227;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN-BERKELEY.
 RX MEDLINE-20196006; PubMed-1073 142;
 KA Adams M.D., Gelniker S.E., Holt R.A., Evans C.A., Gockayne J.D.,
 KA Ananides P.G., Scherer S.E., Li P.W., Hoskins K.A., Gallo R.F.,
 KA George K.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 KA Sutton G.A., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 KA Brundage K.C., Rogers Y.H., Blake J.R., Champ M., Pfeiffer W.D.,
 KA Wang K., Doyle E., Buxton E., Holt G., Nelson C.K., Miklos G.L.G.,
 KA Abril J.F., Ahtajani A., Ai H., Andrews P., Binkley E., Binkley E.,
 KA Ballou K.M., Bass A., Baxendale J., Bayraktarova L., Beasley E.M.,
 KA Benson K.V., Benos P.V., Bertram R.P., Bhattacharya P., Binkley E.,
 KA Borkov N., Botchan M.P., Boyer T., Brinkstein P., Brothier P.,
 KA Burling K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra L.,
 KA Cherry J.M., Chetani S., Clarke S., Davenport L.P., Davies P.,
 KA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew L., Dietz S.M.,
 KA Dodson K., Doup L.E., Downes M., Duran-Koch S., Duncker B.C., Dunn P.,
 KA Durbin K.J., Eganalis C.C., Ferraz C., Ferrera S., Fleischmann W.,
 KA Glodex A., Goni F., Gorrill J.H., Gu Z., Guan P., Harris M.,
 KA Harris N.L., Harvey D., Heimar T.J., Hernandez J.F., Hock J.,
 KA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeigwan C.,
 KA Jalali M., Kalush F., Karpen C.H., Ke Z., Kennison J.A., Ketchum K.A.,
 KA Kimmel B.E., Klotz S., Klotz T., Klotz S., Klotz S., Klotz S.,
 KA Lasko P., Lei Y., Levitsky A., Li J., Li Z., Liang Y., Lin X.,
 KA Liu X., Matten B., McIntosh T.C., McLeod M.P., McPherson D.,
 KA Merkulyov G., Milshina N.V., Murphy B., Morris J., Moshiri A.,
 KA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 KA Nelson D.L., Nelson D.L., Nelson D.L., Nelson D.L., Nelson D.L.,
 KA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 KA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 KA Shier B.C., Shih-Klamon L., Simpson M., Skopets M.P., Smith T.,
 KA Spier E., Spradling A.C., Stapleton M., Strong R., Sun F.,
 KA Svirskas K., Teator C., Turner E., Ventor E., Wand A.H., Wand X.,
 KA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 KA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 KA Ye J., Ye R., Yeh F., Yeh F., Yeh F., Yeh F., Yeh F., Yeh F.,
 KA Zheng X.H., Zhou F.N., Zhou W., Zhou X., Zhou X., Zhou X.,
 KA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 CC -1- SIMILARITY: TO EF-HAND FAMILY.
 DE EMBL: AF003515; AAF49100.2;
 DR HSP: P02588; T1N;
 DR FlyBase: FBgn0046913; CG8434.

Query Match 0.68; Score 64; DB 5; Length 1154;
 Best Local Similarity 29.48; Ref. No. 340-02;
 Matches 20; Conservative 10; Mismatches 48; Indels 0; Gaps 0;
 QY 360 NNIDHILASIFFEYVTSYSSQPFEMISNKVSIINFGELDIPIIFSKYDPTIDVIFEFHLK 449
 DB 1081 NNAVHYMKILPARKTIFISQIMENISFVFIQIDPQKDFEYLKLIVE 1140
 QY 450 ELAULKQ 457
 DB 1141 YVADMFQ 1148
 RESULT 97
 QYVW49
 ID QYVW49 PRELIMINARY PRI: 1779 AA.
 AC QYVW49;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
 DE CG8334 PROTEIN (FRAGMENT).
 GN CG8334.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Erosophilidae; Drosophila.
 CX NCBI_TaxID 7227;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN-BERKELEY.
 RX MEDLINE-20196006; PubMed-1073 142;
 KA Adams M.D., Gelniker S.E., Holt R.A., Evans C.A., Gockayne J.D.,
 KA Ananides P.G., Scherer S.E., Li P.W., Hoskins K.A., Gallo R.F.,
 KA George K.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 KA Sutton G.A., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 KA Brundage K.C., Rogers Y.H., Blake J.R., Champ M., Pfeiffer W.D.,
 KA Wang K., Doyle E., Buxton E., Holt G., Nelson C.K., Miklos G.L.G.,
 KA Abril J.F., Ahtajani A., Ai H., Andrews P., Binkley E., Binkley E.,
 KA Ballou K.M., Bass A., Baxendale J., Bayraktarova L., Beasley E.M.,
 KA Benson K.V., Benos P.V., Bertram R.P., Bhattacharya P., Binkley E.,
 KA Borkov N., Botchan M.P., Boyer T., Brinkstein P., Brothier P.,
 KA Burling K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra L.,
 KA Cherry J.M., Chetani S., Clarke S., Davenport L.P., Davies P.,
 KA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew L., Dietz S.M.,
 KA Dodson K., Doup L.E., Downes M., Duran-Koch S., Duncker B.C., Dunn P.,
 KA Durbin K.J., Eganalis C.C., Ferraz C., Ferrera S., Fleischmann W.,
 KA Glodex A., Goni F., Gorrill J.H., Gu Z., Guan P., Harris M.,
 KA Harris N.L., Harvey D., Heimar T.J., Hernandez J.F., Hock J.,
 KA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeigwan C.,
 KA Jalali M., Kalush F., Karpen C.H., Ke Z., Kennison J.A., Ketchum K.A.,
 KA Kimmel B.E., Klotz S., Klotz T., Klotz S., Klotz S., Klotz S.,
 KA Lasko P., Lei Y., Levitsky A., Li J., Li Z., Liang Y., Lin X.,
 KA Liu X., Matten B., McIntosh T.C., McLeod M.P., McPherson D.,
 KA Merkulyov G., Milshina N.V., Murphy B., Morris J., Moshiri A.,
 KA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 KA Nelson D.L., Nelson D.L., Nelson D.L., Nelson D.L., Nelson D.L.,
 KA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 KA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 KA Shier B.C., Shih-Klamon L., Simpson M., Skopets M.P., Smith T.,
 KA Spier E., Spradling A.C., Stapleton M., Strong R., Sun F.,
 KA Svirskas K., Teator C., Turner E., Ventor E., Wand A.H., Wand X.,
 KA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 KA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 KA Ye J., Ye R., Yeh F., Yeh F., Yeh F., Yeh F., Yeh F., Yeh F.,
 KA Zheng X.H., Zhou F.N., Zhou W., Zhou X., Zhou X., Zhou X.,
 KA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 CC -1- SIMILARITY: TO EF-HAND FAMILY.
 DE EMBL: AF003515; AAF49100.2;
 DR HSP: P02588; T1N;
 DR FlyBase: FBgn0046913; CG8434.

```

DR InterPro: IPR001125; -
DR InterPro: IPR001230; -
DR InterPro: IPR001394; -
DR InterPro: IPR002048; -
DR Pfam: PF00036; ehand; 2.
DR Pfam: PF00442; UCH-1; 1.
DR Pfam: PF00443; UCH-2; 1.
DR PRINTS: PR00450; REG-VERIN.
DR PROSITE: PS00018; EF_HAND; UNKNOWN_1.
DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.
DR PROSITE: PS00972; UCH_2_1; 1.
DR PROSITE: PS00973; UCH_2_2; 1.
DR PROSITE: PS0235; UCH_2_3; 1.
DR Calcium-binding.
KW NON_TER 1
FT SEQUENCE 1779 AA; B0624792A02FA652 CRC64;

Query Match 0.6%; Score 64; DB 5; Length 1779;
Best Local Similarity 36.6%; Pred. No. 5; 02;
Matches 15; Conservative 8; Mismatches 18; Indels 0; Gaps 0;

QY 1230 LRELEPELEPLDQFENRMYVPELLEFVYVQVLESLD 1236
      : : : : : : : : : : : : : : : : : :
DB 268 MINVLIVAKENGSCQYKDLTKQLVISLLEFGQPESPNG 308

RESULT 98
ID 096133 PRELIMINARY; PRT; 1979 AA.
AC 096133;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE HYPOHELICAL 237.7 KUA PROTEIN.
GN PF00145C.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Gardner M.J., Tettelin H., Garucci D.J., Cummings L.M., Atavind L.,
RA Koonin E.V., Shalton S., Masen T., Yu K., Fujii C., Pedison J.,
RA Shen K., Jing J., Aston G., Lai Z., Schwartz D.C., Perle M.,
RA Salberg S., Zhou L., Sutton G., Clayton P., White O., Smith H. O.,
RA Fraser C.M., Adams M.D., Venter J.C., Hoffman S.L.;
RA "Chromosome 2 sequence of the human malaria parasite Plasmodium
RA falciparum".
RT Science 282:1126-1132(1998).
RL EMBL: AF001375; AAC71819.1; -.
KW Hypothetical protein.
SQ SEQUENCE 1979 AA; 237745 MW; 5G6CD6307AFF7D37 CRC64;

Query Match 0.6%; Score 64; DB 5; Length 1979;
Best Local Similarity 21.1%; Pred. No. 6; 2c; 02;
Matches 15; Conservative 24; Mismatches 33; Indels 0; Gaps 0;

QY 387 SLKRLDILLASLFEVYLSVSSGEEESKRVSLNKGFLPLPLESKYPFTLPVWIFE 446
      : : : : : : : : : : : : : : : : : :
DB 766 TLKEVEYK:NTKFEVEYKINTLNEQNHPT:NT:RQNDIKINTMVEEYDQWNTNEQ 825

QY 447 HKLKIADLRKK 457
      : : : : :
DB 826 NEUKMSLKEE 836

RESULT 99
QYQWMS
ID QYQWMS PRELIMINARY; PRT; 136 AA.
AC QYQWMS;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)

```

```

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE NONSTRUCTURAL PROTEIN 1.
GN NS1.
OS Bovine respiratory syncytial virus.
OC Viruses; ssRNA negative strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11246;
RN [1]
RP SEQUENCE FROM N.A.
RA Larsen L.E., Gottschalk E., Blixenkron-Moeller M.;
RA Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DE EMBL: AF054668; AAC36678.1; -.
KW Nonstructural protein.
SQ SEQUENCE 136 AA; 15244 MW; AEB6A8907FADDD24 CRC64;

Query Match 0.6%; Score 63; DB 14; Length 136;
Best Local Similarity 28.1%; Pred. No. 40;
Matches 16; Conservative 12; Mismatches 29; Indels 0; Gaps 0;

QY 161 FWTIPVQSCVPLAKTCTCTHCYKYNCFMDFICSLVTKSVKVFAPVPGSSAOLKVL 217
      : : : : : : : : : : : : : : : : : :
DB 77 FTSMPILONGGYIWMELTHCFQTNGLIDNCEITFSKRLSDSELANTSNQLSILL 133

RESULT 100
ID 077004 PRELIMINARY; PRT; 199 AA.
AC 077004;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE PV200 HIGH-BINDING REGION 1 (FFAGMENT).
OS Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
RP SEQUENCE FROM N.A.
RA Guierrez J.A., Vicini J., Murillo L.A., Patarrayo M.E.;
RA Patarrayo M.A.;
RA Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DE EMBL: U00525; AAC32532.1; -.
FT NON_TER 199
FT NON_TER 199
SQ SEQUENCE 199 AA; 22428 MW; 6CDSA57112F5C748 CRC64;

Query Match 0.6%; Score 63; DB 5; Length 199;
Best Local Similarity 40.9%; Pred. No. 61;
Matches 18; Conservative 2; Mismatches 24; Indels 0; Gaps 0;

QY 1152 KGISVNAEQVRIELEPPDKAKPLGTVOQKRPQKMQKKKSQDLRS 1195
      : : : : : : : : : : : : : : : : : :
DB 28 KGIKALLEQVEAEKKKLPLKONTNTPLTDEQKAAKKIADLES 71

RESULT 101
ID 09TY40 PRELIMINARY; PRT; 199 AA.
AC 09TY40;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE PV200 HIGH-BINDING REGION 1 (FFAGMENT).
OS Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
RP SEQUENCE FROM N.A.
RA Guierrez J.A., Vicini J., Murillo L.A., Patarrayo M.E.;
RA Patarrayo M.A.;

```


A fortiori $Y_{\alpha} \in \mathcal{Y}_{\alpha}$ für $\alpha \in \mathbb{N}$, falls $\mathbb{Z}_{\alpha} \in \mathcal{Z}_{\alpha}$.

OC Pterygota; Neoptera

05 Diptera: *Macropodaphys* (Furn. 17).
06 Eukaryota, Metazoa, Articulata, Tracheata, Insecta;
09 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

DE KINESIN-14K PROTEIN
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID:3702;
RN [1]
RP SEQUENCE FROM N.A.
RA SHINN P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RC Khan S., Kim C., Altati H., Bei B., Chin C., Chou J., Choi E.,
RX Conway A., Gonzalez A., Hansen N., Howing H., Koo T., Lam H.,
MEDLINE:20181125; PubMed 10718197;
RA Sato S., Nakamura Y., Kameo I., Kato T., Asamizu E., Kotani H.,
RA Tabata S.,
RT "Structural analysis of Arabidopsis thaliana chromosome 5, X. Sequence
RT features of the regions of 3,076,755 bp covered by slaty 11 and ZAC
RT clones";
RF LNA Res. 7:31-63(2000).
RF EMBL: AB019225; BAB11107.1;
DR EMBL: AB006707; BAB11107.1; JOINED.
SQ SEQUENCE 967 AA; 108786 MW; 2E03D8016D5A0D15 CRC64.

Query Match 0.6%; Score 63; DB 16; Length 967;
Best Local Similarity 23.9%; Pred. No. 3.7e-02;
Matches 17; Conservative 18; Mismatches 36; Indels 0; Gaps 0;
QY 1145 QTVSSVFKGISVNAQVRILEPPDRAKAPLGTVQKFRGKMKQKKSQGLSVQVGVGYSW 1204
DB 400 ETKHKQLEKLNQAQKAEVLEQVNSFLPVVFAKPLKPLCTKTKPKWKKKQTKYKRF 359
QY 1205 QRVTLLHLEHQA 1215
DB 360 NHOEALQELK 370

RESULT 119
Q9LP05 PRELIMINARY; PRT: 1003 AA.
AC Q9LP05;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE F15H18.12.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID:3702;
RN [1]
RP SEQUENCE FROM N.A.
RA SHINN P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA Kim C., Altati H., Bei B., Chin C., Chou J., Choi E., Conn L.,
RA Conway A., Gonzalez A., Hansen N., Howing H., Koo T., Lam H., Lee J.,
RA Lenz C., Li J., Liu A., Liu J., Lin S., Mukharsky N., Breyer J.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.,
RT "Genomic sequence for Arabidopsis thaliana BAC F15H18 from chromosome
RT 1";
RL Submitted (NOV-1999; to the EMBL/GenBank/DDBJ databases).
EN [2]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
SQ SEQUENCE (NOV-1999; to the EMBL/GenBank/DDBJ databases).
RN [3]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
SQ SEQUENCE (NOV-1999; to the EMBL/GenBank/DDBJ databases).
RN [4]
RP SEQUENCE FROM N.A.
RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
RA Shinn P., Altati H., Bei B., Chin C., Chou J., Choi E., Conn L.,
RA Conway A., Gonzalez A., Hansen N., Howing H., Koo T., Lam H., Lee J.,
RA Lenz C., Li J., Liu A., Liu J., Lin S., Mukharsky N., Breyer J.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.

RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N., Theologis A.,
RA Ecker J.R.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altati H., Bei B., Chin C., Chou J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing H., Koo T., Lam H.,
RA Lee J., Lenz C., Li J., Liu A., Liu J., Lin S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.R.;
EL Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL: AC013354; AAF25984.1;
DR InterPro: IPR001752;
DR Pfam: PF00225; kinesin; 3.
DR PRINTS: PR00340; KINESINHEAVY.
DR PROSITE: PS00667; KINESIN_MOTOR_DOMAIN2; 1.
DR SMART: SM00129; KISC; 1.
KW ATP-binding; Coiled coil; Microtubules; Motor protein.
SQ SEQUENCE 1003 AA; 114706 MW; 3CFHF2596F323B17 CRC64;

Query Match 0.6%; Score 63; DB 10; Length 1003;
Best Local Similarity 48.1%; Pred. No. 3.9e-02;
Matches 13; Conservative 7; Mismatches 7; Indels 0; Gaps 0;
QY 677 PSSMKWEDLISVGEESFNIKQKVT 703
DB 576 PSAMKKEVGVIAFNKSVSANLKEEIT 602

RESULT 119
Q9VZ52 PRELIMINARY; PRT: 1133 AA.
AC Q9VZ52;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CG11203 PROTEIN.
GN CG11203.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA STEVIN REPKELEY;
MEDLINE:20196006; PubMed-10731132;
RA Adams M.D., Gelniker S.F., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananthanades P.G., Scherer S.E., Li P.W., Hoskins R.A., Gallo R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.A.,
RA Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Davis C., Royster J.E., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-pfannkuch C., Baldwin D.,
RA Ballow P.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley R.M.,
RA Beeson A.Y., Berns P.V., Bertram B.F., Blandini E., Bolstakov S.,
RA Borkova D., Botchan M.R., Bourk J., Brokstein P., Brothier P.,
RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cheshy S., Chikhe C., Chivers P., Davies P.,
RA de Fabros B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doop L.E., Dowds M., Duan Pocka S., Dwyer R.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier W.M., Glasser K.,
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glaser K.,
RA Glodok A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris R.L., Harvey D., Helmer T.J., Hernandez J.P., Hovak J.,
RA Jastill D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwa C.,
RA Jallat M., Kalush F., Kaplan G.H., Ke Z., Kennison J.A., Ketchum R.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
RA Lasko J., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

SEQUENCE FROM N.A.
MEDLINE:94150718; PubMed:7906398;
RP Wilson P., Ainsworth R., Anderson K., Baynes C., Berks M.,
RX Penfield J., Burton L., Connell M., Copsey T., Cooper J., Coulson A.,
PA Crakton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
PA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
PA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
PA Lightning J., Lloyd C., Mcmurray A., Mortimore B., O'Callaghan M.,
EA Rausus J., Percy C., Pitke J., Poppra A., Saunders D., Shownkeen R.,
EA Smalton N., Smith A., Sonnenhammer E., Staden K., Sulston J.,
RA Thelery-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wodhiam P.,
RA WT et al. of continuous nucleotide sequence from chromosome III of C.

RA Jones M., Kershaw J., Kirsten D., Jaister N., Latreille P.,
RA Lightning D., Lloyd C., Murray A., Mortimore B., O'Callaghan M.,
RA Tauson L., Tracy K., Pittman L., Roper A., Saunders D., Shewkeen R.,
RA Smalton N., Smith A., Sonnenhammer E., Staden K., Sulston J.,
RA Thiercy-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat F.J., Wohlhadt P.,
RA Wilson R.B. et al. 1987. The complete nucleotide sequence from chromosome III of *C.*

RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome 111 of C.

Oregon: 7
 RL Nature 368:32-38(1994).
 DR EMBL, Z68113; CAA92148.1; -.
 DR HSP; F13569; 1NBD.

```

DR InterPro; TPR003439; -.
DR InterPro; TPR003593; -.
DR Pfam; PF00005; ABC_tran; 2.
DR Pfam; PF00664; ABC_membrane; 2.

```

```

DR PROSITE: PS00211; ABC_TRANSPORTER; HIKKNOWN_1.
DR SMART: SM00382; AAA; 1.
DR SQ SEQUENCE: 1398 AA; 156597 MW; 31C7EDE541ABEDBB CRC64;

Query Match          0.68; Score 63; DB 5; Length 1398;
Pos+ local similarity 29.38; Proj No. 5.6e+02;
Matches 12; Conservative 10; Mismatches 19; Indels 0; Gaps

QY 637 LCRWEEALENVKSTPKCKLIGVANOKMIELLADNINLGD 677

```

Db 428 LSAMETAEETIPVFPKELKMIKQSALLKTFADCI.NVAP 468

RESULT	125
C018191	
ID	C018191
AC	C018191; PRELIMINARY;
DT	01-JAN-1968 (TREMIRE); 05, Created)
ET	01-MAY-1999 (TREMIRE). 10, Last sequence update)
DJ	01-OCT-2000 (TREMIRE). 15, Last annotation update)
DE	WLOC6.1 PROTEIN.

GN W10C6.1.
OS *Caenorhabditis elegans*.
OC Eukaryota; Metazoa; Platyhelminthes; Platyhelminthes; Rhabditida; Rhabditidae;
OC Rhabditidae; Polodrilidae; Caenorhabditis.
OX NCBI_TaxID=6239;
GN

RN [1]
 RP SEQUENCE FROM N. A.
 RA Ainscough R.;

DR	PFam: PF00385; chromo: 1.	DR	PFam: PF00385; chromo: 1.
DR	PROSITE: PS50013; CHROMO_2: 1.	DR	PROSITE: PS50013; CHROMO_2: 1.
KW	ATP-binding; Helicase.	KW	ATP-binding; Helicase.
SQ	SEQUENCE 2703 AA; 295070 MW; 3A5D59DF279E7204 CRC64;	SQ	SEQUENCE 2703 AA; 295070 MW; 3A5D59DF279E7204 CRC64;
Query Match 0.6%; Score 63; DB 5; Length 2703;			
Best Local Similarity 39.5%; Pred. No. 1.2e-03;			
Matches 17; Conservative 7; Mismatches 19; Indels 0; Gaps 0;			
QY	384 LDKSLNNLLHLLASLLEFLVSYCSQELHSEHVELLNEE 424	QY	384 LDKSLNNLLHLLASLLEFLVSYCSQELHSEHVELLNEE 424
DB	2185 LLSGIPQNNISELSLNLFLSQSSVELSECFGSURTER 2227	DB	2185 LLSGIPQNNISELSLNLFLSQSSVELSECFGSURTER 2227
RESULT 129			
Q13719	PRELIMINARY; PRT: 2783 AA.	Q13719	PRELIMINARY; PRT: 2783 AA.
AC	Q13719;	AC	Q13719;
DT	01-NOV-1996 (TREMblrel. 01, Created)	DT	01-NOV-1996 (TREMblrel. 01, Created)
DT	01-NOV-1996 (TREMblrel. 01, Last sequence update)	DT	01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT	01-MAR-2001 (TREMblrel. 16, Last annotation update)	DT	01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE	ALPHA-FETOPROTEIN ENHANCER BINDING PROTEIN.	DE	ALPHA-FETOPROTEIN ENHANCER BINDING PROTEIN.
OS	Homo sapiens (human)	OS	Homo sapiens (human)
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;	OX	NCBI_TaxID=9606;
RN	111	RN	111
RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.
RC	TISSUE=EPIDIDYMA:	RC	TISSUE=EPIDIDYMA:
KX	MEDLINE=92049333; PubMed=1719379;	KX	MEDLINE=92049333; PubMed=1719379;
RA	Morinaga T., Yasuda H., Higashio K., Tamaki T.,	RA	Morinaga T., Yasuda H., Higashio K., Tamaki T.,
RA	"A human alpha-fetoprotein enhancer-binding protein, ATF1, contains	RA	"A human alpha-fetoprotein enhancer-binding protein, ATF1, contains
RT	four homeodomains and seven zinc fingers."	RT	four homeodomains and seven zinc fingers."
RL	Mol. Cell. Biol. 11:6041-6049(1991).	RL	Mol. Cell. Biol. 11:6041-6049(1991).
CC	-1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).	CC	-1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC	-1- SIMILARITY: TO OTHER HOMEBOX DOMAINS.	CC	-1- SIMILARITY: TO OTHER HOMEBOX DOMAINS.
DR	EMBL: D16256; AAA01895.1;	DR	EMBL: D16256; AAA01895.1;
DR	HSSP: P02836; IENH.	DR	HSSP: P02836; IENH.
DR	InterPro: IPR000822;	DR	InterPro: IPR000822;
DR	InterPro: IPR001356;	DR	InterPro: IPR001356;
DR	PFam: PF00046; homeobox; 4.	DR	PFam: PF00046; homeobox; 4.
DR	PFam: PF00096; zf-C2H2; 16.	DR	PFam: PF00096; zf-C2H2; 16.
DR	PROSITE: PS00027; HOMEBOX.1; 2.	DR	PROSITE: PS00027; HOMEBOX.1; 2.
DR	PROSITE: PS00071; HOMEBOX.2; 4.	DR	PROSITE: PS00071; HOMEBOX.2; 4.
DR	PROSITE: PS00026; ZINC_FINGER_C2H2; 11.	DR	PROSITE: PS00026; ZINC_FINGER_C2H2; 11.
TP	SMART: SM00389; HOX; 1.	TP	SMART: SM00389; HOX; 1.
KW	DNA-binding; Homeobox; Metal-binding; Nuclear protein; Zinc-finger.	KW	DNA-binding; Homeobox; Metal-binding; Nuclear protein; Zinc-finger.
FT	VARIANT 2460 2460 A -> V.	FT	VARIANT 2460 2460 A -> V.
SQ	SEQUENCE 2783 AA; 305747 MW; 623A59DF279E7204 CRC64;	SQ	SEQUENCE 2783 AA; 305747 MW; 623A59DF279E7204 CRC64;
Query Match 0.6%; Score 63; DB 4; Length 2783;			
Best Local Similarity 27.9%; Pred. No. 1.2e-03;			
Matches 19; Conservative 13; Mismatches 36; Indels 0; Gaps 0;			
QY	1178 GGRGPKMKKKSKQLPSVQKGGSYWQVRVILLELLHKKKKRSFQIVPTFLNL.SRC 1237	QY	1178 GGRGPKMKKKSKQLPSVQKGGSYWQVRVILLELLHKKKKRSFQIVPTFLNL.SRC 1237
DB	810 GGRGPKMKKKSKQLPSVQKGGSYWQVRVILLELLHKKKKRSFQIVPTFLNL.SRC 1237	DB	810 GGRGPKMKKKSKQLPSVQKGGSYWQVRVILLELLHKKKKRSFQIVPTFLNL.SRC 1237
QY	1238 LFPDPEQ 1245	QY	1238 LFPDPEQ 1245
DB	870 TPTTIQLQ 877	DB	870 TPTTIQLQ 877
RESULT 130			
Q9NDJ2	PRELIMINARY; PRT: 3201 AA.	Q9NDJ2	PRELIMINARY; PRT: 3201 AA.
AC	Q9NDJ2;	AC	Q9NDJ2;
DT	01-OCT-2000 (TREMblrel. 15, Created)	DT	01-OCT-2000 (TREMblrel. 15, Created)
DT	01-OCT-2000 (TREMblrel. 15, Last sequence update)	DT	01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT	01-MAR-2001 (TREMblrel. 16, Last annotation update)	DT	01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE	HELICASE DOMINO A.	DE	HELICASE DOMINO A.


```

; UNKNOWN_1.
7061;7061;2 1464;

```

```

; DB 5; Length 5107;
; 2.5e+03;
atches 37; Indels 0; Gaps 0;

```

KLVQPPFFPFPPFPAIINPDAITTKL 1677
:
EVNVAQNQVGFPPKQRIEAKIKSEDL 391

170 AA.

```

source update)
tation update)
.
every Phabditida: Phabditoidae;
tis.

```

Barnes C., Barks M.,
 Bussey T., Cooper J., Coulson A.,
 Favely T., Fulton I.,
 Hier I., Tier M., Johnston L.,
 Hester N., Latreille P.,
 Martimore R., McCallaghan M.,
 Mearns A., Saunders D., Showackeen R.,
 Staden R., Sulston J.,
 Vaughan K., Waterston R.,
 Weir D., Wohlhuth P.;
 sequence from chromosome III of C.

ark/DBD databases.

ank/DBJ databases.
DEBFAE03A7B CRC64;

2.5e03;	Indels	0;	Gaps	0;
atches	19;			
KGSDPSVQFVC	1200			
EATDQEAETLC	4795			

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh K.F., Zaccari J.S., Zhao M., Zhang C., Zhao Q., Zhong L.,
 RA Zhou X.H., Zhou E.H., Zhou K.W., Zhou S., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:
 RA "The genome sequence of *Drosophila melanogaster*."
 SC Science 287:2185-2195(2000).
 DR EMBL: AF003734; AAF55892.1;
 DR FlyBase: FB003846; CG6475.
 DR InterPro: IPR002213;
 DR Pfam: PF00201; UDEGT_1.
 DR PROSITE: PS00375; UDEGT_1.
 SQ SEQUENCE 441 AA; 50073 MW; AR74BA044714C0B2 CRC64;

Query Match 0.6%; Score 62; DB 5; Length 441;

Best Local Similarity 30.4%; Pred. No. 26+02;

Matches 11; Conservative 15; Mismatches 29; Indels 0; Gaps 0;

QY 374 TKEHLEADLEKLSERNHGLASLPHYVLSVSGPHEKVKGLRQFNP 427

DB 141 VIEFMFSTIEFVTEVFQITSVLITSEMTITQFVWQNMVSWC;HTLP 194

RESULT 158

QVW759

ID QVW759 PRELIMINARY; PRT; 443 AA.

QVW759

DT 01-NOV-1999 (TEMBLrel. 12, Created)

DT 01-NOV-1999 (TEMBLrel. 12, Last sequence update)

DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)

DE 94.6% IDENTICAL TO US GENE OF STRAIN U1102 OF HHV-6.

OS Human herpesvirus 6.

OC Viruses, dsDNA viruses, no RNA stage; Herpesviridae;

OC Betaherpesvirinae; Roseolovirus.

OX NCBI_TaxID=10368;

QY SEQUENCE FROM N.A.

RA Isebara Y., Makai T., Nakano K., Nakae M., Chan J., Mori Y.,

RA Saitawa T., Sashihara J., Kato T., Kosaka H., Yamanishi K.,

RA "A comparison of the complete DNA sequences between human herpesvirus

6, variant A and B."

RA J. Virol. 0:0-0(1999).

DR EMBL: AB021506; BAA78225.1; -.

SQ SEQUENCE 443 AA; 51515 MW; 0D1F168070355191 CRC64;

Query Match 0.6%; Score 62; DB 14; Length 443;

Best Local Similarity 30.0%; Pred. No. 2,10+02;

Matches 18; Conservative 9; Mismatches 33; Indels 0; Gaps 0;

QY 1392 EHRPLPLVMTLGAERTWILLILFFQYVTKTVLAAYGKDAILEDTEFWFVC 1451

DB 188 EHSRPLSLTSTVTPDKKALLTLR;VFIVFVVT;YKVINAEGLIYYDLEFEVGEKLC 247

RESULT 159

QVW759

ID QVW759 PRELIMINARY; PRT; 451 AA.

QVW759

DT 01-MAY-2000 (TEMBLrel. 13, Created)

DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)

DT 01-JUN-2000 (TEMBLrel. 14, Last annotation update)

DE CG12960 PROTEIN.

OS *Drosophila melanogaster* (fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; *Drosophila*.

OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 PX MFR1 INF-20196306, PubMed 10731132;
 RA Adams M.D., Celivner S.E., Holt P.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George P.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yeung M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H.C., Blakes B.G., Champ M., Pfeiffer B.D.,
 RA Wan K.H., Dewis C., Rafter P.C., Holt C., Nelson C.P., Milos C.L.G.,
 RA Abrell J.F., Aghayani A., An H.-J., Andrews-Trankoch C., Baldwin D.,
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Bottlier P.,
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra L.,
 RA Chert J.M., Cawley S., Dahlke C., Davenport I.R., Davies P.,
 RA de Pablos R., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Drenth J., Doup E.E., Downes M., Duan Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.T., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glaser A., Gong P., Corvelli J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Heston D., Houston K.A., Howland T.J., Wei M.-H., Ibeqwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel R.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
 RA Lasko F., Lei Y., Levitsky A.A., Li J., Li Y., Liang Y., Lin X.,
 RA Liu X., Mattei R., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mohanty C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paciej J.M.,
 RA Palamara M., Pittman C.S., Pan S., Pollard T., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamus I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Switskas F., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh K.F., Zaccari J.S., Zhao M., Zhang C., Zhao Q., Zhong L.,
 RA Zhou X.H., Zhou E.H., Zhou K.W., Zhou S., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:
 RT "The genome sequence of *Drosophila melanogaster*."
 SC Science 287:2185-2195(2000).
 DP EMBL: AF003734; AAF55892.1;
 DE FlyBase: FB003846; CG12960.
 SQ SEQUENCE 491 AA; 57573 MW; 8CB320ABBE95A0F5 CRC64;

Query Match 0.6%; Score 62; DB 5; Length 491;

Best Local Similarity 28.8%; Pred. No. 2,3e+02;

Matches 16; Conservative 9; Mismatches 31; Indels 0; Gaps 0;

QY 1900 LLLQFLNCLYKFLFDTQPTSKERAAALMPLVQLENLGGGERKQKVKHL 2045

DB 203 LSTFVCLFSLVLIIVTQHLNLSKNTLANILLNDRKSLRGLGSGFPFPPNPKHL 258

RESULT 160

Q90407

ID Q90407 PRELIMINARY; PRT; 546 AA.

Q90407

DT 01-MAY-2000 (TEMBLrel. 13, Created)

DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)

DT 01-OCT-2000 (TEMBLrel. 15, Last annotation update)

DE RED-BIOLUMINESCENCE ELICITING LUCIFERASE.

OS *Phrixothrix hirtus*.

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;

OC Elateriformia; Cantharidae; Phengodidae; Phrixothrix.

OX NCBI_TaxID=94779;

QY SEQUENCE FROM N.A.

PP MEDLINE=99315203; PubMed=10387072;

PX

[illegible]

72301 PRELIMINARY; 649 AA.
15 p72301
C p72301;

Matches	13;	Conservative	13;	Mismatches	21;	Indels	0;	Gaps	0;
---------	-----	--------------	-----	------------	-----	--------	----	------	----


```

GN MIRA OR MIRANDA OR CG12249
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=CANTON-S;
RX MEDLINE=98065952; PubMed=9403694;
RA Ikeshima-Kataoka H., Skeath J.B., Nabeshima Y., Doe C.Q.,
RA Matsuzaki F.;
RT "Miranda directs Prospero to a daughter cell during Drosophila
PT asymmetric divisions.";
RE Nature 390:625-629(1997);
DR EMBL: AB005661; BAA2411.1; -.
DR FlyBase: FBgn0021776; miranda.
SQ SEQUENCE 830 AA; 93171 MW; F7AD159AA58F178A CPG64;

Query Match 0.6%; Score 62; DB 5; Length 830;
Best Local Similarity 45.7%; Pred. No. 4.2e+02;
Matches 14; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

Q7 1176 TVGKPPKMKQKKSQLESVQEVGGSYWQ 1205
      |||  |||  |||  |||  |||
Db 402 TEQQAQAVKVVQVESDLESAREKSYWR 431

RESULT 171
O45116 PRELIMINARY; PRT: 830 AA.
AC O45116;
DT 01-JUN-1998 (TRENBLREL 06, Created)
DT 01-JUN-1998 (TRENBLREL 06, Last sequence update)
DT 01-MAR-2001 (TRENBLREL 16, Last annotation update)
DE MIRANDA.
GN MIRA OR MIRANDA OR CG12249.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=CANTON-S;
RX MEDLINE=97410310; PubMed=9267025;
RA Shen C.P., Jan L.Y., Jan Y.N.;
RT "Miranda is required for the asymmetric localization of Prospero
RT during mitosis in Drosophila.";
RL Cell 90:449-458(1997).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=CANTON-S;
RA Shen C.P., Jan L.Y., Jan Y.N.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
DR EMBL: AF045771; AAC02621.1; -.
DR FlyBase: FBgn0021776; miranda.
SQ SEQUENCE 830 AA; 93172 MW; 7JF88E8144F41134 CPG64,

Query Match 0.6%; Score 62; DB 5; Length 830;
Best Local Similarity 46.7%; Pred. No. 4.2e+02;
Matches 14; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

Q7 1176 TVGKPPKMKQKKSQLESVQEVGGSYWQ 1205
      |||  |||  |||  |||  |||
Db 402 TEQQAQAVKVVQVESDLESAREKSYWR 431

RESULT 172
O44224 PRELIMINARY; PRT: 836 AA.
AC O44224;
DT 01-JUN-1998 (TRENBLREL 06, Created)
DT 01-JUN-1998 (TRENBLREL 06, Last sequence update)
DT 01-MAR-2001 (TRENBLREL 16, Last annotation update)
DE MIRANDA.
GN MIRA OR MIRANDA OR CG12249.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

```



```

DI 01-MAY-2000 (TrEMBLrel. 15, last sequence update)
DE GENMIP1A, CHROMOSOME 5, BAC clone pF66.
OS Arabidopsis thaliana (Mouse ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Malvales; Euphorbiales; Euphorbiaceae; Rosales; eucrocodius II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_taxID=375;
LN [1]
RN SEQUENCE FROM N.A.
RP STRAIN S-UMBBIA;
RA Katoke I., Katoke I., Asanuma Y., Katoke H.,
RA Ito H. S.;
RI "Structural analysis of Arabidopsis thaliana chromosome 5, XI."
RI Submitted (2011-09-09) to the EMBL/GenBank/CCDS databases.
ID EMBL: AF557966; BAC clone pF66;
SQ SEQUENCE 896 AA; 94361 MW; A7e71204A3704EB C6064;

Query Match. 0.00; Score 62; DB 10; Length 896;
Best Local Similarity 41.00; Pred. No. 4,2e+02;
Matches 13; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Q7 644 EMBL:AF557966:EMBL:AF557966:EMBL:AF557966:701
Q8 111 111 111 111 111 111 111 111 111 111
Q9 4 4 KVLKKGKNSLGMALAAAGKGLLLEELSLNQLQ 442

RESULT 174
Q2004 2004 PRELIMINARY; PRI: 963 AA.
A 2004
Q2004 2004 (TrEMBLrel. 14, created)
Q2004 01-MAY-2000 (TrEMBLrel. 14, last sequence update)
DI 01-MAY-2000 (TrEMBLrel. 14, last sequence update)
DE 01-MAY-2000 (TrEMBLrel. 14, last sequence update)
ID STRAIN 229; COMPLETE GENOME.
LN [1]
RN Human herpesvirus 6B.
RP Virus; dsDNA viruses, no RNA stage; Herpesviridae;
OC Borna disease virus; Borna disease virus.
OX NCBI_taxID=4043;
LN [1]
RN SEQUENCE FROM N.A.
RP STRAIN 229;
RI METLINE 99412418; PubMed 10482553;
RA Tomlinson G., Lambaugh T.R., Stanley F.R., Dewhurst S.,
RA Pellett P.E.;
RI "Human herpesvirus 6B genome sequence: coding content and comparison
RI with human herpesvirus 6A."
RI J. Virol. 73:6040-6052(1999).
RN SEQUENCE FROM N.A.
RP STRAIN 229;
RA Pellett P.E., Tomlinson G., Lambaugh T.R., Stanley F.R., Dewhurst S.,
RA Tomlinson G.;
RI Submitted (2011-09-09) to the EMBL/GenBank/CCDS databases.
ID EMBL: AF557966; AAL4 64317;
SQ SEQUENCE 613 AA; 104221 MW; A618E6430E6E312 C6064;

Query Match. 0.00; Score 62; DB 14; Length 908;
Best Local Similarity 30.00; Pred. No. 4,6e+02;
Matches 13; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Q7 149 EMBL:AF557966:EMBL:AF557966:EMBL:AF557966:1451
Q8 111 111 111 111 111 111 111 111 111 111
Q9 4 4 EMBL:AF557966:EMBL:AF557966:EMBL:AF557966:707

RESULT 174
P4044
ID P4044 PRELIMINARY; PRI: 985 AA.
A P4044
Q1 01-MAY-2000 (TrEMBLrel. 00, created)

```

```

DI 01-MAY-1997 (TrEMBLrel. 03, last sequence update)
DE 01-MAY-2000 (TrEMBLrel. 13, last annotation update)
DE ALPHA-KETOGLUTARATE DEHYDROGENASE.
GN SUCA.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; alpha subphylum; Rhizobiales; alpha;
OC Bradyrhizobium group; Bradyrhizobium.
OX NCBI_taxID=375;
LN [1]
RN SEQUENCE FROM N.A.
RP STRAIN USDA110;
RI MEDLINE-97136621; PubMed 858198;
RA Green L.S., Emerich D.W.;
RI "Bradyrhizobium japonicum does not require alpha-ketoglutarate
RI dehydrogenase for growth on succinate or malate."
RI J. Bacteriol. 179:194-201(1997).
ID EMBL: U74618; AAC44748.1;
DR InterPro: IPR001017;
DR Pfam: PF00676; E1-dehydrog; 1.
SQ SEQUENCE 965 AA; 110920 MW; C4D4E9C48E44280 C6064;

Query Match. 0.00; Score 62; DB 2; Length 985;
Best Local Similarity 43.48; Pred. No. 5,1e+02;
Matches 13; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Q7 1774 EMBL:U74618:EMBL:U74618:EMBL:U74618:1802
Q8 111 111 111 111 111 111 111 111 111 111
Q9 4 4 HTEPYLEWLVNQVNSVRRRYVGRASA 962

RESULT 175
Q900W2 PRELIMINARY; PRI: 1046 AA.
ID Q900W2 Q24002; Q90182;
AC Q900W2; Q24002; Q90182;
DI 01-MAY-2000 (TrEMBLrel. 13, created)
DI 01-MAY-2000 (TrEMBLrel. 16, last sequence update)
DI 01-MAY-2000 (TrEMBLrel. 16, last sequence update)
DE TKR PROTEIN (TKR-III PROTEIN).
LN [1]
RN Drosophila melanogaster (Fruit fly).
OS Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_taxID=7227;
LN [1]
RN SEQUENCE FROM N.A.
RP STRAIN BERKELEY;
RI METLINE-20196006; PubMed 10731132;
RA Adams M.D., Gelniker S.E., Holt R.A., Evans C.A., Gockyne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.F., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brannon R.C., Rogers Y.H.C., Hlazo R.G., Champ M., Pfeiffer R.D.,
RA Wan F., Joyly R., Bayrer F., Jasti A., Nelson C.W., Myers G.L.G.,
RA April J.F., Aghayani A., An H.J., Andrews-Plunkett C., Baldwin L.,
RA Ballou K.M., Basa A., Baxendale J., Kayaktaradali L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,
RA Chertys K.C., Husam D.A., Butler H., Cadion E., Center A., Chandra L.,
RA Curtis J.M., Cawley S., Dahike C., Davenport L.H., Davies P.,
RA de Fabios R., Delcher A., Deng Z., Dey S.A.D., Dew L., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dusan-Rocha S., Dunkey B.C., Dunn P.,
RA Durbin K., Evangelista C., Ferraz C., Forster S., Fleischmann W.,
RA Foster C., Gabrielson A.E., Garb N.S., Gelbart W.M., Glasser K.,
RA Glisler A., Gung F., Gurell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.H., Harvey D., Heiman T.J., Hernandez J.R., Hock J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwa C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.B., Li J., Li Z., Liang Y., Lin X.,
RA Lin X., Martel B., McIntosh T.C., Melov M.P., McPherson D.,
RA Merkulov G., Milshina N.V., McBarry C., Morris J., Moshrefi A.,

```


Fri Nov 9 15:42:42 2001

DR EMBL: AF054665; AAC36669.1; -;
 DR EMBL: AF054664; AAC36666.1; -;
 KW Nonstructural protein.
 SQ SEQUENCE 136 AA; 15258 MW; AEB6A89D78FCD24 CRC64;

Query Match 0.6%; Score 61; DB 14; Length 136;
 Best Local Similarity 28.1%; Pred. No. 72;
 Matches 16; Conservative 12; Mismatches 20; Indels 0; Gaps 0;

QY 161 FMLPVPKSGVGLAKGILITITVFELFPMFICISLVTRSVKVFAPYFGSSAGRVLL 217
 DB 77 FTSMPILONGGYWELMELDFEQTNCLIDNCEITFSKPLSLSLAKYSNQLSILL 133

RESULT 182
 ID Q9L559 PRELIMINARY; PRT; 163 AA.
 AC Q9L559;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE GLINOLINIC ACID SYNTHETASE (FRAGMENT).
 OS NADA.
 OC Pseudomonas fluorescens.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OC NCBI_TaxID=294;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-267;
 RC Marek Kowczuk M., Skorupska A M.
 RT "Acquisition of a gene involved in biosynthesis of niacin in
 Pseudomonas fluorescens strain 267.";
 RL Submitted (APR-2000) to the EMBL/Genbank/Trna databases.
 FM EMBL: AF257101; AAF72624.1; -;
 FT NON_TER 1
 SQ SEQUENCE 163 AA; 17620 MW; 085E1C1A255B6F0 CRC64;

Query Match 0.6%; Score 61; DB 2; Length 163;
 Best Local Similarity 40.7%; Pred. No. 89;
 Matches 11; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 497 LAMHLKKIMKTSKEGVDESEIKRAVL 523
 DB 105 MAMULERTLSKEGTNEIFVDVALI 131

RESULT 183
 ID Q9FOA0 PRELIMINARY; PRT; 200 AA.
 AC Q9FOA0;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE GLUTATHIONE S-TRANSFERASE GST 34 (EC 2.5.1.18)
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade, Panicoideae;
 OC Andropogoneae; Zea.
 OC NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE: 20532786; PubMed=11080289;
 RA McGonigle B., Koeler S.J., Iau S.M.C., Koeppe M.K., O'Keefe D.P.;
 RT "A Genomics Approach to the Comprehensive Analysis of the Glutathione
 S-Transferase Gene Family in Soybean and Maize";
 RL Plant Physiol. 124:1105-1120(2000).
 DP EMBL: AF244704; AAG4847.1; -;
 KW Transferase.
 SQ SEQUENCE 200 AA; 23127 MW; 666C34072E147DB9 CRC64;

Query Match 0.6%; Score 61; DB 10; Length 200;
 Best Local Similarity 40.0%; Pred. No. 110+02;
 Matches 12; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 1094 ELYAGMPTIQTALFKITKFFFAAISDEKV 1123
 DB 78 ELFAGRPILPTDPEYATATAPFAAYADDKL 107

RESULT 184
 ID Q28135 PRELIMINARY; PRT; 237 AA.
 AC Q28135;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE PROLACTIN-LIKE PROTEIN PRECURSOR.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA OF 6 MONTH GESTATION;
 RC MEDLINE-91198142; PubMed 2015300;
 RA Tanaka M., Minoura H., Gshiro H., Nakashima K.;
 RT "A novel cDNA clone encoding a prolactin like protein that lacks the
 two C-terminal cysteine residues isolated from bovine placenta.";
 RL Biochim. Biophys. Acta 1088:385-389(1991).
 DR EMBL: X59504; CAA42092.1; -;
 DR HSSP: Q2H642; IAN3.
 DR InterPro: IPR001400; -;
 DR Pfam: PF00103; hormone; 1.
 DR PRINTS: PR00836; SOMATOTROPIN.
 KW Signal.
 KW SIGNAL 1 36 POTENTIAL.
 FT CHAIN 37 237 PROLACTIN-LIKE PROTEIN.
 SQ SEQUENCE 237 AA; 27377 MW; E3B703F17E18BA3 CRC64;

Query Match 0.6%; Score 61; DB 6; Length 247;
 Best Local Similarity 39.5%; Pred. No. 140+02;
 Matches 15; Conservative 3; Mismatches 20; Indels 0; Gaps 0;

QY 1255 TSLCLNLCQKLSPEGKIPKGTITDFKFRVFLVQCI 1292
 DB 31 LLLCGNSCPSCCPDVSDIPUDLLRELFNLATLSOSI 68

RESULT 185
 ID Q9PJG4 PRELIMINARY; PRT; 246 AA.
 AC Q9PJG4;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE TYPE III SECRETION CHAPERONE SYCD.
 OS Chlamydia muridarum.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OC NCBI_TaxID=83560;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MORN / NIGG;
 RA MEDLINE 20150255; PubMed-10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
 RA Linher K., Weidman J., Khouri H., Graves B., Bowman C., Rodson R.,
 RA Selim M., Reison W., DeBoy E., Feltz J., McCarthy G., Galzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RL "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia

DR EMBL: Z81074; CAB03036.1;
 SSO SEQUENCE 278 AA; 36942 MW; 98D1E19E501F862E CRC64;
 Query Match 0.6%; Score 61; DB 5; Length 278;
 Best Local Similarity 22.4%; Pred. No. 156; 02;
 Matches 15; Conservative 13; Mismatches 39; Indels 0; Gaps 0;
 QY 250 PEEAATYMTPTTSVPTRENTVTHVLAASITPTPTFLIEGICSTLVLDLQPKES 315
 DB 23 EPEETVPTCTVETAKPTALNFVSTFWTTTATATATPWSISPTATGAPDIF 82
 QY 315 LKKEKP 322
 DB 83 VTSPVP 89
 RESULT 169
 QY 456 PRELIMINARY; FRT: 296 AA.
 AC Q93456;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE F1028.7 PROTEIN.
 GN F1028.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Phabditida; Phabditidae;
 OC Rhabditidae; Poloderinae; Caenorhabditis.
 NCBI_TaxID=6239;
 OX 111
 RP SEQUENCE FROM N.A.
 RA Basham V.;
 RN Submitted (SEF 1996) to the EMBL/GenBank/DBJ databases.
 RL [2]
 QY MEDLINE:94150738; PubMed 7905498;
 RX Wilson R., Ainsworth K., Anderson K., Baynes C., Becks M.,
 Bentley J., Burton J., Censell M., Gopsey T., Cooper J., Coulson A.,
 Craxton M., Dear S., De Z., Durbin K., Eaves A., Fulton L.,
 Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 Jones M., Korshaw J., Kingston J., Laister N., Latreille P.,
 Lightning J., Lloyd C., Murray A., Mortimore B., O'Callaghan M.,
 Parsons J., Percy C., Riken L., Raza A., Saunders B., Shawcross P.,
 Smalton N., Smith A., Souhammer E., Staden R., Sulston J.,
 Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 Watson A., Weinstock L., Wilkinson S., Winton S., Winton J.,
 RT Nature 368:32-38 (1993).
 RL Nature 362:16; CAB02283.1;
 DR EMBL: Z80216; CAB02283.1;
 DR InterPro: IPR000445;
 DR SEQUENCE 296 AA; 43614 MW; A1E1594221280BA CRC64;

Query Match 0.6%; Score 61; DB 5; Length 296;
 Best Local Similarity 26.7%; Pred. No. 136; 02;
 Matches 12; Conservative 10; Mismatches 24; Indels 0; Gaps 0;
 QY 1522 PFWVSTZPHLFLGHEFLGVGVAVAGQEFETATVST 1576
 DB 210 FFAVLAAGVSGVSESESEFMAVSEETAAPEETETAWVLEP 25;
 RESULT 190
 QY 600 PRELIMINARY; FRT: 298 AA.
 AC Q96000;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE D1HYR3B11GCOLLATE SYNTHETASE.
 GN D1HYR3B11GCOLLATE SYNTHETASE.
 OS D1HYR3B11GCOLLATE SYNTHETASE.
 OC D1HYR3B11GCOLLATE SYNTHETASE.
 NCBI_TaxID=1;
 OX 1;
 RP SEQUENCE FROM N.A.
 RA Smith T.;
 RN Submitted (SEF 1996) to the EMBL/GenBank/DBJ databases.
 RL [2]
 QY MEDLINE:94150738; PubMed 7905498;
 RX Wilson R., Ainsworth K., Anderson K., Baynes C., Becks M.,
 Bentley J., Burton J., Censell M., Gopsey T., Cooper J., Coulson A.,
 Craxton M., Dear S., De Z., Durbin K., Eaves A., Fulton L.,
 Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 Jones M., Korshaw J., Kingston J., Laister N., Latreille P.,
 Lightning J., Lloyd C., Murray A., Mortimore B., O'Callaghan M.,
 Parsons J., Percy C., Riken L., Raza A., Saunders B., Shawcross P.,
 Smalton N., Smith A., Souhammer E., Staden R., Sulston J.,
 Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 Watson A., Weinstock L., Wilkinson S., Winton S., Winton J.,
 RT Nature 368:32-38 (1993).
 RL Nature 362:16; CAB02283.1;
 DR EMBL: Z80216; CAB02283.1;
 DR InterPro: IPR000445;
 DR SEQUENCE 296 AA; 43614 MW; A1E1594221280BA CRC64;

OS Zymomonas mobilis.
 OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonas group;
 CC Zymomonas.
 NCBI_TaxID=542;
 RN SPOUNFP FROM N.A.
 PC STRAIN-ZM4;
 RA Lee J.S., Kang H.S.;
 RT "S. pom. mobilis 347 44171" found close of zymomonas mobilis ZM4";
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 DE EMBL: AF117451; AAF17453.1;
 DR InterPro: IPR002220;
 DR Pfam: PF00761; DHDPS; 1;
 DR PRINTS: PR00146; DHPICSNTHASE.
 DR SEQUENCE 299 AA; 33180 MW; 3A12C938653FE11 CRC64;
 Query Match 0.6%; Score 61; DB 2; Length 298;
 Best Local Similarity 28.5%; Pred. No. 140; 02;
 Matches 16; Conservative 15; Mismatches 25; Indels 0; Gaps 0;
 QY 657 TGVANQKMIELLADNINLQDPSSMLKMWEDLISVGFPSFNLPQVTFRVIIIV 712
 DB 76 TGAFTPTDILLSEQLQSGVSALLPLVSYHALNDEVSLEYTVSKHVSVIV 131
 RESULT 191
 QY 600 PRELIMINARY; FRT: 322 AA.
 AC Q9VD19;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE CARNITIC AMYLASE (EC 4.2.1.1) (CAPRONATE DEHYDRATASE)
 GN CG5379.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE 20190036; PubMed 10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gabyone J.D.,
 Amaratunga C., Scherer S.E., Li P.W., Hoskins P.A., Gall P.E.,
 George R.A., Lewis S.F., Richards S., Ashburner M., Henderson S.N.,
 Sutton G.G., Warren S.R., Yandell M.P., Zhang Q., Chen L.X.,
 Brandon P.C., Rogers Y.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter K.G., Holt R., Nelson C.P., Miklos G.L.G.,
 Abell T.F., Aghayani A., An H.-T., Andrews P., Anukool C., Baldwin D.,
 Ballwey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Benson K.Y., Berns P.V., Berube R.P., Boudouri F., Borkov S.,
 Borkova D., Botchan M.R., Boudreau N., Brockstein P., Brothier P.,
 Burtis K.C., Busam D.A., Butler H., Caden E., Center A., Chandra L.,
 Cherry J.M., Cawley S., Dahlke C., Davenport J., Davies P.,
 De Pablos P., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dusen K., Durr I.E., Dwan M., Egan J., Fierman J.C., Dunn P.,
 Durbin K.J., Evangelista C., Ferraz C., Ferrier S., Fleischmann W.,
 EA Fisher R., Gabrielian A.E., Gars H.S., Gilbert W.M., Glasser K.,
 Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris H.L., Harvey D., Heiman T., Hernandez J.P., Hoare J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam C.,
 Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Pulp P., Tai Z.,
 Lasko F., Lei Y., Levitsky A.A., Li J., Lin X., Lin X., Lin X.,
 Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 Merkley C., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson P.P., Nelson K.A., Nixon K., Nusskern P.R., Pacle J.M.,
 Palazzone M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reinert S., Remington K., Saunders F.P.C., Scheeler P., Shen H.,
 Shure B.C., Sidon-Kiamos I., Simpson M., Skupski M.P., Smith T.,

PA Shieh E., Spradling A.C., Stapleton M., Strong P., Sun R., Sun R.,
 PA Schirakus R., Tector A., Turner K., Walter E., Ward A.B., Wang X.,
 PA Wang Z.Y., Watanabe D.A., Weinstock G.M., Willey K., Wu D., Yang S., Yao Q.A.,
 PA Williams S.M., Woodard T., Worley K.F., Wu D., Yang S., Yao Q.A.,
 PA Yeh J., Yeh K.F., Zeng J., Zeng J.S., Zhan M., Zhang G., Zhao G., Zhou L.,
 PA Zhou X.H., Zhou F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 PA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 PT "The genome sequence of *Drosophila melanogaster*.";
 PL Science 287:2246-2254 (2000).
 SC 1. FUNCTION: REVERSIBLE HYDRATATION OF 'APRON DIOXIDE' (BY
 SC SIMILARITY).
 SC 2. CATALYTIC ACTIVITY: H₂O₂(aq) + O₂(g) + H₂O (BY SIMILARITY).
 SC 3. MISCELLANEOUS: THERE ARE AT LEAST 7 ENZYMATIC FORMS OF CARRIER
 SC ANHYDRASE: CA-I (OR C), CA-II (OR C), CA-III (OR M), CA-IV, CA-V,
 SC CA-VI AND CA-VII (BY SIMILARITY).
 SC 4. SIMILARITY: TO THE EUKARYOTIC TYPE CARRIER ANHYDRASE FAMILY.
 LP EMBL: A000334; AAF55964.1; 5.
 LP HSPD: Q64344; 228.
 LP FlyBase: FB0009696; 059479.
 LP InterPro: IPR001148; 1.
 LP Pfam: PF00194; carbanhydrase; 2.
 LP ProDom: PD00065; 1.
 LP PROSITE: PS00162; F0K_D02; ANHYDRASE; 1.
 KW Tracheal gland.
 SQ SEQUENCE 422 AA: 4720 MW: 10784707FA2DF619 GR064;

Query Match 0.64; Score 61; DB 10; Length 422;
 Best local similarity 49.44; Pred. No. 20-02;
 Matches 11; Conservative 5; Mismatches 13; Indels 0; Gaps 0;
 97 1750 LSALAAKQKVFTHFISYEGTISQVHL 1789
 100 164 VIVLAVLFKVKVPSDFEYQIGSELSNHL 214

RESULT 192
 Q9XIN? 0.64; Score 61; DB 10; Length 445;
 ID Q9XIN? PRELIMINARY; PRI: 445 AA.
 AC Q9XIN?;
 DI 01 NOV 1999 (EMBLrel. 12, Last sequence update)
 DI 01 MAR 2001 (EMBLrel. 16, Last sequence update)
 DE POTATIVE NAM PROTEIN.
 GN F1426.12.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Malvaceae; Malvaceae; Malvaceae; core eudicots; Rosidae; eudicots II;
 OC Brassicaceae; Brassicaceae; Arabidopsis.
 CX NCBI_TaxID:3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN:CV, COLUMBIA;
 RA Liu X., Kuhl S., Shen L.D., Fritschy J.N., Shen M., Vazirani S.P.,
 RA Barnstead M.L., Mison F.M., Bowman C.L., Bonini C.M., Bonito M.,
 RA Bartlett A.L., Crosby J.H., Boffi G.R., Town C.D., Nierman W.C.,
 RA Fraser C.M., Venter J.C.,
 RA "Arabidopsis thaliana chromosome 11 BAC F1426 genomic sequence";
 RC Submitted (02-1999) to the EMBL/GenBank/DBJ databases.
 DE EMBL: A000253; AAF1199.1; 5.
 LP InterPro: IPR00441; 5.
 LP Pfam: PF0205; NAM; 1.
 SQ SEQUENCE 445 AA: 38184 MW: 8084107F616134E GR064;

Query Match 0.64; Score 61; DB 10; Length 445;
 Best local similarity 49.44; Pred. No. 20-02;
 Matches 11; Conservative 6; Mismatches 11; Indels 0; Gaps 0;
 97 169 FACKTEWLRHREHLYNQSLIAVVL 146
 100 129 PRGRFWIMHLYTHADQALVAVVQL 154

RESULT 193
 Q9M9I2 0.64; Score 61; DB 10; Length 439 AA.
 ID Q9M9I2; PRELIMINARY; PRI: 439 AA.
 AC Q9M9I2;
 DI 01-OCT-2000 (EMBLrel. 15, Created)
 DI 01-OCT-2000 (EMBLrel. 15, Last sequence update)
 DI 01-OCT-2000 (EMBLrel. 15, Last annotation update)
 DE F1417.6 protein.
 GN F1417.6.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Malvaceae; Malvaceae; Malvaceae; core eudicots; Rosidae; eudicots II;
 OC Brassicaceae; Brassicaceae; Arabidopsis.
 CX NCBI_TaxID:3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN:CV, COLUMBIA;
 RA Liu X., Kuhl S., Shen L.D., Fritschy J.N., Shen M., Vazirani S.P.,
 RA Barnstead M.L., Mison F.M., Bowman C.L., Bonini C.M., Bonito M.,
 RA Bartlett A.L., Crosby J.H., Boffi G.R., Town C.D., Nierman W.C.,
 RA Fraser C.M., Venter J.C.,
 RA "Arabidopsis thaliana chromosome 11 BAC F1417 genomic sequence";
 RC Submitted (02-1999) to the EMBL/GenBank/DBJ databases.
 DE EMBL: A000253; AAF1199.1; 5.
 LP InterPro: IPR00441; 5.
 LP Pfam: PF0205; NAM; 1.
 SQ SEQUENCE 439 AA: 37685 MW: 675AEE017AF6E07 GR064;

Query Match 0.64; Score 61; DB 10; Length 439;
 Best local similarity 49.58; Pred. No. 20-02;
 Matches 15; Conservative 7; Mismatches 16; Indels 0; Gaps 0;
 97 550 SEVITSNINIPQAFISFNGFWIVIVIRIAAHLIRKE 587
 100 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 12 62 SEFAAALVNSQKILAFKMLQALIKVAMIMIRPE 99
 RESULT 194
 Q9G454 0.64; Score 61; DB 10; Length 341 AA.
 ID Q9G454; PRELIMINARY; PRI: 341 AA.
 AC Q9G454;
 DI 01 MAR 2001 (EMBLrel. 15, Created)
 DI 01 MAR 2001 (EMBLrel. 16, Last sequence update)
 DE NADH DEHYDROGENASE SUBUNIT 2.
 GN ND2.
 OS Traptelus sanduolentus.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidodactylia; Squamata; Iguania; Acrodonta; Agamidae; Agamidae;
 OC Traptelus.
 CX NCBI_TaxID 118238;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE-07154820; PubMed-9000751;
 RA Macey J.R., Larson A., Ananjeva N.H., Papenfuss T.J.;

Best Local Similarity 30.2%, Pred. No. 2,20,02;
Matches 13; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

QY 1472 KLPEKEETIPKAVSPNFSQSPFMGVENVETHUSKQIPHEK 1514
||||||| : || : ||| : |||| :
DB 172 KLPGGGGQSPGQIVSAPEKPKPEPPQIKQGLEDMQKQIRQLQ 214

Search Completed. December 6, 2001, 12:24:51
Job time: 278 sec

Complete version 4.5
Copyright (c) 1993 - 2000 CompuLink Ltd.

OM protein: protein search using SW model

Run date: November 6, 2001, 12:16:44 : Search time: 63.5 seconds

(without dist. cuts)

4733,862 Million cell updates/sec

Database: US-09-603-665-5

Protein score: 10807

Sequence: 1 MSV194 (P1A1P0582)

CONFIDENTIAL: CERN-SWF 2144

Scoring table: 16, 8000

Gapop factor: 0.0001

Searches: 219241 seqs, 76174662 residues

Total number of hits satisfying chosen parameters: 219241

Minimum hit seq length: 1

Maximum hit seq length: 250000

Post processing: Minimum March 08

Maximum March 13

Database: PIR, 681*

1: p1r1*

2: p1r2*

3: p1r3*

4: p1r4*

Procl. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	IL	Description
1	239	2.2	176	2	853376	probable membrane
2	199	1.8	169	2	159608	hypothetical prote
3	166	1.5	165	2	127684	hypothetical prote
4	76	0.7	1626	2	14692	neuropilin - tiss
5	74	0.7	1431	2	864761	visc protein - ye
6	74	0.7	408	2	107648	hypothetical prote
7	74	0.7	463	2	10678	desmin - chicken
8	73	0.7	489	2	10678	desmin - human
9	73	0.7	490	2	871279	probable serine-ty
10	73	0.7	1942	2	A19204	translation clon
11	73	0.7	1127	2	128417	ORF MSV196 hypothe
12	72	0.7	153	2	102139	calcium-dependent
13	72	0.7	1369	2	139211	calc protein - Adr
14	71	0.7	1307	2	130887	146b nuclear prote
15	71	0.7	1322	2	124140	hypothetical prote
16	71	0.7	1296	2	843467	RNA-directed RNA p
17	70	0.6	209	2	10678	desmin - human
18	70	0.6	369	2	152409	conserved hypotet
19	70	0.6	469	2	152409	desmin - rat
20	70	0.6	469	2	A51104	desmin - mouse
21	70	0.6	469	2	A51104	desmin - mouse
22	69	0.6	521	2	864541	phosphoprotein, phe
23	69	0.6	749	2	A96742	unknown protein F5
24	69	0.6	242	2	171872	conserved hypotet
25	68	0.6	237	2	864514	hypothetical prote
26	68	0.6	286	2	135152	hypothetical prote
27	68	0.6	409	2	128455	ORF MSV194, A11 not
28	68	0.6	454	2	10678	two-component sens
29	68	0.6	456	2	A12629	regulatory protein

MRH1 protein, limpo	737	2	850942
DNA mismatch repair	737	2	151620
protein F12K11.7	819	2	886201
MAIS box protein 1	246	2	17023
MAIS box protein A	248	2	819534
MAIS box protein A	262	2	751409
desmin - African e	458	2	A15554
hypothetical prote	486	2	146483
hypothetical prote	1159	2	15963
hypothetical prote	68	2	107208
MAIS box protein A	250	2	819534
protein tyrosine k	503	1	1VMS807
limb deformity (ld	1213	2	A17234
hypothetical prote	1231	2	124415
probable Hsp70 prot	1316	2	750545
lps receptor, Xls	2693	2	A10743
hypothetical prote	113	2	G12566
MAIS box protein c	233	2	110714
hypothetical prote	286	1	A11090
methyl-accepting c	433	2	864120
hypothetical prote	941	2	S09851
hypothetical prote	1013	2	146422
translation clon	1044	1	1VBYE3
probable membrane	1044	2	S02926
hypothetical prote	1217	2	110607
dna-directed RNA p	1372	2	817224
hypothetical prote	226	2	853777
methionyl-tRNA for	305	2	861426
glucan endo 1,3 b	343	2	B18257
FA489.7 protein	400	2	S44810
NLS-related prote	416	2	832231
vincentin - chinea	438	2	148128
probable hmo p, ate	450	2	671145
vincentin - golden	464	1	VHHY
vincentin - mouse	466	2	A25074
vincentin - rat	466	2	A43803
dnak-type molecule	605	2	149837
probable repli val	721	2	671409
hypothetical prote	789	2	883949
c-11-ORF phosphol	790	2	175372
conserved hypotet	971	2	170128
p51 protein - ye	1089	2	S63978
hypothetical prote	1622	2	671622
toxic acid synthas	2504	1	A57789
probable lic prote	211	2	870563
type 11 site-speci	247	2	S06035
hypothetical prote	352	2	124279
involved in spec	522	2	A83801
head-to-tail join	535	2	S07521
ankyrin-related pr	834	2	105282
probable membrane	971	2	S54595
replication licens	1017	1	S64219
hypothetical prote	1189	2	151491
isolectin - FNA 1	1266	2	159314
DNA-directed RNA p	1316	2	S41146
hypothetical prote	1398	2	129434
probable protein 3	1447	2	686474
hypothetical prote	1461	2	126327
alpha-tetraprotein	2784	1	A51948
partial vhs - can	5107	2	124144
hypothetical prote	5178	2	115348
auxin-induced prot	75	2	105706
hypothetical prote	174	2	863979
probable low conc	242	2	871497
probable kkt-type	243	2	S52649
probable ribosomal	245	2	872465
M79 protein - rice	249	2	131307
glycoprotein al pr	250	2	382165
enoyl-acyl carrier	275	2	871964
Abp-ATP carrier pr	301	2	S51132
hypothetical prote	364	2	123121
conserved hypotet	470	1	873341

[illegible]

10
11

A:Cross: [refseq=25](#) GR_NC_061268; NID:[q7267357](#); PIDN:[CAR78038.1](#); CSPDP:[GM001440](#)
C:Genetics:
A:Gene: [AT4q09140](#)
A:Map position: 4
C:Superfamily: mismatch repair protein MLH1

Query Match:	0.6%	Score 6.9;	DB 2;	Length 737;
Best Local Similarity	43.8%	Pred. No. 27;		
Matches 14;	Conservative	5;	Mismatches 13;	Indels 0;
Gaps	0;			

Q7 116B PPKAKPLCTVAGQPPKAMQCKKSQDLESQVEV 1199
||| ||| ||| ||| ||| ||| ||| ||| |||
db 409 PPKVSSI SVVPSSSVROPNPKNPTADISSVOEL 460

31
RESULT
T51620

DNA mismatch repair protein MCH1 [imported] - Arabidopsis thaliana (Arabidopsis)
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 03-Nov-2000
 C:Accession: U51620
 R:Jean, M.; Pelletier, J., Hilpert, M.; Reizakis, E.; Kunze, R.
 Mol Gen Genet 262, 633-642, 1999

A;Title: Isolation and characterization of *AtMLH1*, a MutL homologue from Arabidopsis
A;Reference number: Z25418

A;Accession: M51620
A;Status: preliminary; translated from CB/EMBL/DDBJ
A;Molecule type: mRNA

A;Residues: 1 737 <JEA>
A;Molecule type: mon
A;Cross-references: EMBL:AA012747; PDB:CAA10163.1

C;Genetics:
A;Gene: MLI1

A;Map position: 4
C;Superfamily: mismatch repair protein MLH1

Query Match	Score 68;	DB 2;	Length 737;
-------------	-----------	-------	-------------

Best Local Similarity 43.8%; Pred. No. 27;
Matches 14; Conservative 5; Mismatches 13; Indels 0; Gaps 0

Q7 1168 PDKAKPIGTVOOKPPROKMOOKKSQDLESVOEV 1199
||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 429 PDKVSSLSVPPSSVRQPPNPKEITADLSVQEL 460

RESULT 32

protein F12K11.7 [imported] - Arabidopsis thaliana C86201

C:Species: Anabropsis chariana (mouse ear cress)
C:Date: 02 Mar 2001 #Sequence_revision: 03 Mar 2001 #rev_chapno: 41-Mar-2001
C:Accession: CR6201

P.; Theologis, A.; Ecker, J. F.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Chin, C.W.; Chaud, M.K.; Conn, L.; Conway, A.P.; Creasy, T.H.

ansen, N.F.; Hughes, R.; Huizar, L.
Nature 408, 816-820, 2000

A: Authors; Hunter, J.B.; Jenkins, G.G.; Johnson-Nopoul, G.; Kuo, S.-C.; Li, C.-C.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.-X.; Liu, Z.-A.; Luross, J.S.; Maiti, P.; Piao, M.; Rooney, T.; Powley, D.; Sakano, H.

Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sutt, H. Ker, M., Wu, D.; Yu, G.; Frusci, C.M.; Venter, J.C.; Davis, R.W.

A>Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A:Reference number: A86141; MUID:21016719

A: Accession: 196201
A: Status: preliminary
A: Molecule type: DNA

A;Residues: 1 819 <STO>
A;Cross references: GB:AE005172, MID:96692677; PIDN:AAV24811.1; GSPDB:GN00141

C:G4orf11.5.
A:Gene: F12K11.7

A:Map position: 1

1. The first step in the process is to identify the problem or issue that needs to be addressed. This involves gathering information and understanding the context of the problem.

Q: 2125 Cc:142...
DB: 242 P...
RESULT 47
T26483
hypothetical protein Y1707B.1 Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T26483
R:Kershaw, J. submitted to the EMBL Data Library, January 1998
A:Reference number: 225222
A:Accession: T26483
A:Status: preliminary; translated from GR/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1466 SWILL
A:Cross-references: EMBL:AL021472, F108,CAA16299.1, G3F3B,GN03023, GESP:Y1707B.1
A:Experimental source: clone Y1707B
A:Genetics:
A:Map position: 5
A:Introns: 42...
Query Match 0.6%; Score 67; DB 2; Length 486;
Best Local Similarity 23.9%; Pred. No. 23;
Matches 16; Conservative 14; Mismatches 16; Indels 0; Gaps 0;
QY 833 FEMMLGACAGVHFVKMK...
DB 270 FENVGSEK...
RESULT 38
T15963
hypothetical protein F07F6.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15963
A:Status: preliminary; translated from GR/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-1159 <CHI>
A:Cross-references: EMBL:022486; NID:4716417, F108,AA:46777.1, GESP:F07F6.4
A:Experimental source: strain Bristol N2
A:Genetics:
A:Gene: CHSP:F07F6.4
A:Introns: 27...
Query Match 0.6%; Score 57; DB 2; Length 1159;
Best Local Similarity 21.6%; Pred. No. 63;
Matches 22; Conservative 20; Mismatches 50; Indels 0; Gaps 0;
QY 949 ASPEYLLDHLISKAREITSDAAVYVDDIATLPEELQREKKIKSHOKISLTLEKLLSCVY 1008
DB 200 SAVSVYVYKPTKATLCAKFNALGAL...
QY 1009 SCPSYIAKMLKMLGVNGMLVSLQIPMAFQAKTKQKPT 1050
DB 260 ELAKKKSL...
RESULT 49
T07208
hypothetical protein 66 - Chlorocella vulgaris chloroplast

C:Species: chloroplast Chlorocella vulgaris
C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 21-May-2000
C:Accession: T07208
P:Wakasugi, T.; Nagai, T.; Kapoor, M.; Saito, M.; Ito, M.; Hata, S.; Tsudzuki, J.; Na
Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997
A:Title: Complete nucleotide sequence of the chloroplast genome from the green alga C
A:Reference number: 215985; NCID:97303241
A:Accession: T07208
A:Status: preliminary; translated from GR/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-68 <WAK>
A:Cross-references: EMBL:AR001084, NID:3222452, F108,BAAC5855.1, F108-4020471
A:Genetics:
A:Genome: chloroplast
A:Keywords: chloroplast
Query Match 0.6%; Score 66; DB 2; Length 68;
Best Local Similarity 44.8%; Pred. No. 31;
Matches 13; Conservative 5; Mismatches 11; Indels 0; Gaps 0;
QY 69 LPSGLAKTLEPSVCTKAVNKKLNDENISLF 97
DB 1 MESLQKFLIKSVTKTKTKHHTISNLELF 29
RESULT 40
D39534
MAOS box protein AG14 - Arabidopsis thaliana
R:Altman, J.; floral homeotic protein: AG14
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 24-Sep-1999
C:Accession: D39534
R:Ma, H.; Yanofsky, M.F.; Meyerowitz, E.M.
Genes Dev. 5, 484-495, 1991
A:Title: AG14, an Arabidopsis gene family with similarity to floral homeotic and
A:Reference number: A39534, MIM:91160981
A:Accession: D39534
A:Molecule type: DNA
A:Residues: 1-253 <MAA>
A:Cross-references: DB M55553, NID:4165592, F108,AAA3734.1, F108-4165594
C:Genetics:
A:Gene: AG14
C:Superfamily: Transcription factor super, Serum response factor DNA binding domain ho
C:Keywords: DNA binding; nucleus; transcription regulation
F:2-57/Domain: Serum response factor DNA-binding domain homology <SRF>
Query Match 0.6%; Score 66; DB 2; Length 250;
Best Local Similarity 19.3%; Pred. No. 14;
Matches 16; Conservative 25; Mismatches 42; Indels 0; Gaps 0;
QY 1828 ETKQLEKNNKNNHMPMTTQHTGCMKKKEITSHQSQTAPFLALDFRACHSENDE 1887
DB 93 EFYKKEGTENLQRCFHLGCEP...
QY 1888 EVSKTEHCILDLVAMVVKISFV 1910
DB 153 DLOCKEHHILLDANRALSMKLEDM 175
RESULT 41
TVM5HC
Protein tyrosine kinase (EC 2.7.1.112) hck - mouse
N:Alternate names: Kinase-related transforming protein (bmk)
C:Species: Mus musculus (house mouse)
C:Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 29-Jan-2000
C:Accession: A27282; A39973
R:Klemsz, M.J.; Mckercher, S.R.; Makl, R.A.
Nucleic Acids Res. 15, 9600, 1987
A:Title: Nucleotide sequence of the mouse hck gene.
A:Reference number: A27282; MIM:88067781
A:Accession: A27282

C.Species: Sarcharomyces cerevisiae
C.Date: 27-Apr-1996 #sequence_revision: 03-May 1996 #text_change: 02-Feb-2001
C.Accession: S62926
R.Andre, B.; Itakqi Houssaini, I.; Urrestarazu, L.A.; Vissers, S.
Submitted to the Protein Sequence Database, April 1996
A.Reference number: S62920
A.Accession: S62926
A.Molecule type: DNA
A.Residues: 1-1044 <ANI>
A.Cross references: EMBL:J71966, NID:J331925, FIDG:CA56874.1, PIR:J331926, MIPS:YNI01
A.Experimental source: strain S288C
A.Genetics:

A: Map position: 14L
 C: Superfamily: 14L
 G: Superfamily: translation elongation factor 3, ATP binding cassette homology
 K: Keywords: ATP; nucleotide binding; P-loop; transmembrane protein
 L: C keywords: ATP; nucleotide binding; P-loop; transmembrane protein
 F: 248-251/69415 transmembrane status predicted 1982
 F: 358-374/69415 transmembrane status predicted 1982
 F: 416_616/69416 transmembrane status predicted 1982
 F: 463-470/region: nucleotide-binding motif A (P-loop)
 F: 664_664/69416 transmembrane status predicted 1982
 F: 701-708/region: nucleotide-binding motif A (P-loop)

Query Match.	0.6%	Score 45.	98.2%	Length 1044;
Best Local Similarity	43.38%	Pred. No. 1e+02;		
Matches	17;	Conservative	8;	Mismatches 26, Indels 6, Gaps 3,
QY	2079	EVELEAAITVLAALAEI	FFRYVILFEVFIPLAEIMLITFEFFVHDCQPT	3129
b_144	144	KPAVAVLEIVSVLVAAAGAEAL	PMPELLELVESPMVTEKKGVKAATITTI	194

RESULT 54
100607
hypothetical protein A2292556 (unpublished). Arabidopsis thaliana
N/A:Alternate names: Hypothetical protein TBK22.11
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 01-feb-1999 #sequence_revision 01-feb-1999 #text_change 02 Feb 2001
C:Accession: T00607; B84438
R:Rausley, S.D., Lin, X., Kelcham, K.A., Crosby, M.L., Braddon, E.C., Sykes, S.M., Kaul
submitted to the EMBL Data Library, June 1998
A:Description: Arabidopsis thaliana chromosome II BAC TBK22 genomic sequence.
A:Reference number: Z14192
A:Accession: T00657
A:Status: translated from G5/EMBL/LDB3
A:Molecule type: DNA
A:Molecule type: RBC>
A:Residues: 1 1217 #EOC>
A:Cross references: EMBL,ACC04136, N14-93184273, F1D-9318183
A:Experimental source: Cultivar Columbia
R:Lin, X.; Kaul, S.; Koutsuley, S.D.; Shew, T.P.; Reulke, M.T.; Town, C.D.; Fujii, C.Y.;
M.M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, J.;
Euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
This sequence was analyzed as part of chromosome 2 of the plant Arabidopsis thaliana

A:Accession: B84438
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1217 <ST>
A:Cross references: GR:AF920942; NIDM-AACT14030; 1- CSPJPR-GN00139
C:Genetics:
A:Gene: A72q02560; 19822.14
A:Map position: 2
A:Mutations: 142/3, 562/2, 148/3, 194/4, 239/2, 275/3, 322/2, 380/3, 411/2, 443/3, 4

	Query Match	0.68	Score 55	DB 23	Length 12177
	Best Local Similarity	29.84	Prod. No. 1.26e92		
	Matches	17	Mismatches	31	Gaps 0
QY	1688	FFVVLVATVAVLAFAPKEKKEKVLGSLDGLAEVTSLEALAIPLQSLMPSLLTMMK	1746		
UB	554	FFVIVVIVVAVINERITRQGGCVFVCAITLMELVSTHCTGEPAEFLTSCLAVLVQVQW	612		

A:Title: Nucleotide sequence of a Bacillus megaterium gene homologous to the dnaK gene
A:Reference number: 139837; MUID:87231083
A:Accession: 139837
A:Status: preliminary; translated from GR/UMK/01/01/01
A:Molecule type: DNA
A:Residues: 1-605 <RES>
A:Cross-references: GB:M31338; NID:q39628; PID:CAA6348.1; PID:q39629
C:Function: involved in protein folding and assembling/disassembling of protein complex
C:Superfamily: heat shock protein 70
C:Keywords: ATP; molecular chaperone

Query Match 0.6%; Score 64; DB 2; Length 605;
Best Local Similarity 32.7%; Pred. No. 73; Indels 0; Gaps 0;
Matches 17; Conservative 12; Mismatches 23

QY 548 FSSSEVTISRLNIPQAPISYNNPWFYVIFIAAPETIKETISEPQISNOV 599
DB 514 FTTEKTLKDLGKVEAEVTPANFADAKAATFNINMPPKAKKPTQPIV 565

RESULT 68
G71409
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thaliana
A:Reference number: A71400; MUID:98121113
A:Accession: G71409
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-771 <REV>
A:Cross-references: GB:297336; EID:q2244788; PID:q2244826
C:Genetics:
A:Map position: 4COP9-4G3845

Query Match 0.6%; Score 64; DB 2; Length 771;
Best Local Similarity 36.1%; Pred. No. 97; Indels 0; Gaps 0;
Matches 13; Conservative 10; Mismatches 13

QY 639 GWEPATFNVIKSTKPKKIGVANKMIFETLAPNTNL 674
DB 487 GWKALGSLNEPFAEGRKIGRENERKPCILLDELGV 522

RESULT 69
C83949
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83949; MUID:20263314
A:Accession: C83949
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-789 <STO>
A:Cross-references: GB:A0001515; GB:BA000004; NID:q10174886; PID:BAH06114; PID:GN00
A:Experimental source: strain C-125
C:Genetics:

Query Match 0.6%; Score 64; DB 2; Length 789;
Best Local Similarity 33.3%; Pred. No. 99; Indels 0; Gaps 0;
Matches 13; Conservative 8; Mismatches 18

QY 1470 LKLPKEKETIPKAVSPNKSEQEMLOVFNVEHHSK 1508
DB 323 LKLPNNPQSMKRLILKNAKRLKRLTLESFGVKAHVK 361

RESULT 70
F75372
A:Title: GMP phosphodiesterase A - Deinococcus radiodurans (strain R1)
A:Reference number: A75372; MUID:20036896
A:Accession: F75372
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-799 <WHI>
A:Cross-references: GB:A1302007; EID:q456443; PID:AAH11199; PID:q445
A:Experimental source: strain R1
C:Genetics:
A:Map position: 1

Query Match 0.6%; Score 64; DB 2; Length 790;
Best Local Similarity 34.0%; Pred. No. 1e+02; Indels 0; Gaps 0;
Matches 17; Conservative 9; Mismatches 24

QY 1190 SQPLESVQGVGGVWQVILLLELLQHKKLLKSPQILVPTLNLISRCLE 1239
DB 297 SLNACVQLSGSGTGAATITTFPALPRAKTVASPRYKSVLPFHSQILLD 256

RESULT 71
D70128
A:Title: Conserved hypothetical protein B60228 - Lyme disease spirochete
A:Reference number: A70128; MUID:98065943
A:Accession: D70128
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-971 <YLF>
A:Cross-references: GB:A0001515; GB:BA000004; NID:q456443; PID:AAH11199; PID:q445
A:Experimental source: strain B31

Query Match 0.6%; Score 64; DB 2; Length 971;
Best Local Similarity 34.1%; Pred. No. 1.3e+02; Indels 0; Gaps 0;
Matches 14; Conservative 9; Mismatches 19

QY 61 SFEQEPAPLPSLAKTLERSVQTKAVNKDENTISLILHL 101
DB 839 SFASVDPNFTKTYQAFKSLLELANKNMTDDEITYLIGL 879

A:Gene: GIB:ILRS
A:Cross-references: GDB:384085; OMIM:600709
A:Map position: 9q21-q21
C:Superfamily: Isoleucine--Lysine
C:Keywords: Lysine

Query Match 0.6%; Score 63; DB 2; Length 1266;
Best Local Similarity 28.3%; Pred. No. 2.3e+02;
Matches 15; Conservative 10; Mismatches 28; Indels 0; Gaps 0;

QY 295 PLEDTYLRKRGAFVLMRVVKELESTGVVQCATTCAGAEVEMDNILIR 347
DB 295 PLEDTYLRKRGAFVLMRVVKELESTGVVQCATTCAGAEVEMDNILIR 347

RESULT 85

S31146
DRA directed RNA polymerase (D2 2.7.7.6) beta' chain - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 30 Sep 1993 #sequence_revision 30 Sep 1993 #text_change 18-Aug-1993
C:Accession: S31146
R:Honore, N.; Bergh, S.; Chanteau, S.; Doucet-Populaire, F.; Eiglmeyer, K.; Garnier, T.; del, P.R.; Slittsombut, N.; Wu-Hunter, S.; Cole, S.T.
Mol. Microbiol. 7, 207-214, 1993
A:Title: Nucleotide sequence of the first cosmid from the Mycobacterium leprae genome project
A:Reference number: S31143; MIMD:93188701
A:Accession: S31146
A:Status: nucleic acid sequence not shown, translation not shown
A:Molecule type: DNA
A:Residues: 1-1316 <NON>
A:Cross-references: EMBL:S31141, DDB:914182, FIDN:CAA78669.1; PDB:55P3ase
A:Notes: the nucleotide sequence was submitted to the EMBL Data Library, August 1993
C:Genetics:
A:Gene: rpoC
A:Start codon: GIG
C:Superfamily: Escherichia coli DNA directed RNA polymerase beta' chain
C:Keywords: nucleotidyltransferase, transcription

Query Match 0.6%; Score 63; DB 2; Length 1316;
Best Local Similarity 33.3%; Pred. No. 2.4e+02;
Matches 10; Conservative 13; Mismatches 7; Indels 0; Gaps 0;

QY 1523 GLLSRRRLFAVWVSESEFEILFIFFEITIF 1562
DB 446 KVINRSLKRLDLCGAFDILVRNKRMLQ 375

RESULT 86

T20434
Hypothetical protein E9322.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15 Oct-1999 #text_change 18-Feb 2000
C:Accession: T20434
R:McMurray, A.
Submitted to the EMBL Data Library, November 1995
A:Reference number: Z19275
A:Accession: T20434
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1398 <WIL>
A:Cross-references: EMBL:Z048115, FIDN:CAA92348.1; GSPDB:GN09028; GESP:E0902.2
A:Experimental source: clone E03G2
C:Genetics:
A:Gene: GESP:E03G2.2
A:Map position: X
A:Introns: 735-1195, 1275-1595, 1765-1815, 1845-1875, 1975-2043
C:Superfamily: human multidrug resistance protein ABCA2; ATP-binding cassette hemology

Query Match 0.6%; Score 63; DB 2; Length 1398;
Best Local Similarity 29.3%; Pred. No. 2.6e+02;

Matches 12; Conservative 10; Mismatches 19; Indels 0; Gaps 0;
QY 637 LKWEAEALNVTSTKPKGLGVANCKMIELLADRNILGDP 677
DB 428 LSAWETAMEETIERVEDKELMKIKQSALLKTFADTNVAP 468

RESULT 87

G86474
pro-ubiquitin protein-ubiquitin-tyrosine (ubiquitin) - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02 Mar-2001 #sequence_revision 02 Mar-1996 #text_change 31-Mar-2001
C:Accession: G86474
P:Theologis, A.; Ecker, J.P.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon, Chiu, C.W.; Chung, M.K.; Gant, L.; Gaudy, A.B.; Conway, A.E.; Gross, T.H.; Dewar, Ansen, N.P.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Zhai, Y.; Kim, C.A.; Li, H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Maiti, R.; Marzita, Piro, M.; Poon, T.; Poyly, D.; Sakano, H.
A:Accession: S31146
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A96141; MIMD:21016719
A:Accession: G86474
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1447 <STO>
A:Cross-references: GR AE005172, NID:q11034946; FIDN:ANG37103.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 0.6%; Score 63; DB 2; Length 1447;
Best Local Similarity 28.9%; Pred. No. 2.7e+02;
Matches 14; Conservative 14; Mismatches 22; Indels 0; Gaps 0;

QY 351 LPILVSIHHVTCFETFMGGIYKHHHEATITKISKNHILIASLL 400
DB 352 MPQISVNAVSGISGKYTKTVKGVIVKPNLFIHSGSTHNPINSTVAALK 401

RESULT 88

T26327
Hypothetical protein W1006.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T26327
R:Ainscough, R.
Submitted to the EMBL Data Library, September 1997
A:Reference number: Z20199
A:Accession: T26327
A:Status: preliminary; translated from G1/EMBL/700B
A:Molecule type: DNA
A:Residues: 1-1461 <WIL>
A:Cross-references: EMBL:Z99269; FIDN:CAB16467.1; GSPDB:GN00020; GESP:W1006.1
A:Experimental source: clone W1006
C:Genetics:
A:Gene: GESP:W1006.1
A:Map position: ?
A:Introns: 1-1673, 1762-1842, 1900-1963, 1985-2043, 2075-2127, 2155-2193

Query Match 0.6%; Score 63; DB 2; Length 1461;
Best Local Similarity 29.4%; Pred. No. 2.8e+02;
Matches 14; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 1879 AGHFGGFFVYKTFNCTIDGLVAVWKLSEVT 1911
DB 136 AETTRKLLERKHEPCVVDVDAAFVVKSGIT 168

RESULT 89

RESULT 104

C:Species: *Drosophila melanogaster*
C:Date: 31-Dec-1991 #sequence_revision: 31 Dec 1991 #text_change: 22-Jun-1999
C:Accession: S00381; A38801
C:Bossy, B.; Hallivet, M.; Spierer, P.
EMBO J. 7, 611-618, 1988
A:Title: Conservation of neutral nicotinic acetylcholine receptors from *Drosophila* to
A:Reference number: S00381; M81D:88283626
A:Accession: S00381
A:Molecule type: DNA
A:Residues: 1 567 <ROS>
A:Cross-reference: GR_X07194; MID:q7575; PIDN:CAA30172.1; PID:q7576
A:Accession: A38A01
A:Molecule type: mRNA
A:Residues: 1 567 <ROS>
A:Cross-reference: EMBL:X07194; MID:q7575; PIDN:CAA30172.1; PID:q7576
A:Note: 538 Tyr was also found

A: Cross-references: FlyBase:FBgn0000036
A: Ref position: ap 96A
A: Introns: 64/83, 79/83, 116/2, 177/3, 309/2, 401/1, 453/3
C: Subfamily: 307/3, 316/1b, 316/1c
C: Keywords: glycoprotein; ion channel; neurotransmitter receptor; postsynaptic membrane
F: 1-27/Domain: signal sequence #status predicted <SIG>
F: 27-67/Domain: extracellular #status predicted <EX>
F: 67-230/Domain: transmembrane #status predicted <TM>
F: 230-264/Domain: transmembrane #status predicted <TM>
F: 264-290/Domain: transmembrane #status predicted <TM>
F: 290-306/Domain: transmembrane #status predicted <TM>
F: 306-325/Domain: transmembrane #status predicted <TM>

C-Accession: A93987, A93484, S42524; A03900
 C-Strainway: G. G. Stanley, P. H. Mountford, P. C. Keene, P. J. Minor, P. F. Schild, G. C. A. S.
 Proc. Natl. Acad. Sci. U.S.A. 81, 1539-1543, 1984
 A-Title: Comparison of the complete nucleotide sequences of the genomes of the neurov
 A-Reference number: A93987; MUID:84170338
 A-Accession: A93987
 A-Molecule type: genomic RNA
 A-Residues: 1 5206, 5811
 A-Cross-references: GB:K01392, NID:932997, FIDR:AA06014.1; PID:932896
 A-Experimental source: strain Sabín vaccine P3/León/37
 A-Note: The strain Sabín vaccine P3/León/37 is the progenitor of the strain Sabín vac
 A-Note: the authors translated the codon GAG for residue 457 as Gly
 P-Strainway: G. G. Cann, A. J. Hauptmann, R. J. Hughes, P. J. Clarke, L. D. J. Mountford, R. C. J.
 Nucleic Acids Res. 11, 5629-5643, 1983
 A-Title: The nucleotide sequence of poliovirus type 3 loc 12 a 1b, comparison with po
 A-Reference number: A93484, MUID:83299239
 A-Accession: A93484
 A-Molecule type: genomic RNA
 A-Residues: 1 426, 57, 442-853, 17, 865-967, 17, 999-2206, 6213
 A-Cross-references: GR X00925, GB:K00032, NID:961154, FIDR:CAA25444.1; PID:q61155
 A-Experimental source: strain Sabín vaccine P3/León 12a111b
 A-Note: The authors translated the codon GAG for residue 457 as Gly
 P-Strainway: G. G. Cann, A. J. Hauptmann, P. J. Clarke, L. D. J. Mountford, R. C. J.
 Eur. J. Biochem. 135, 529-533, 1983
 A-Title: Nucleic acid sequence of the region of the genome encoding capsid protein VP

A: Molecule type: genomic RNA
A: Residues: 579 878 1578A
A: Cross-references: EMBL: V01407; N101961.5; 1; LDR: CAA24780.1; F1D: q39811
A: Experimental source: strain Nabin vaccine P4/1990/37
A: Summary: Coat proteins VP2 and VP3 and the RNA-directed RNA polymerase are related
C: Superfamily: poliovirus genome polypeptide
C: Keywords: genome linked protein; nucleosyl transferase; phosphoprotein
F: 2-69/Product: coat protein VP4 #status predicted <VP4>
F: 70-340/Product: coat protein VP2 #status predicted <VP2>
F: 341-578/Product: coat protein VP3 #status predicted <VP3>
F: 579-878/Product: coat protein VP1 #status predicted <VP1>
F: 879-1453/Product: core protein P2.3b #status predicted <P2.3>
F: 1027-1453/Product: core protein P2.5b #status predicted <P2.5>
F: 1124-1453/Product: core protein P2.X #status predicted <P2.X>
F: 1541-2206/Product: protein P3.1b #status predicted <P3.1>
F: 1541-1562/Product: genome linked protein VPg #status predicted <Vg>
F: 1563-2206/Product: protein P3.2 #status predicted <P3.2>
F: 1563-2206/Product: protein P3.3 #status predicted <P3.3>
F: 1563-1745/Product: probable proteinase P3.7c #status predicted <P3.7>
F: 1746-2206/Product: RNA-directed RNA polymerase P3-4b #status predicted <P3.4>
F: 1543/Binding site: phosphoryl-RNA (5'p) (covalent) #status predicted

A:Authors: Ferreira, V.C.A.; Porto, J.A.; Prata, J.S.; Franca, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kasper, T.B.; Kijakajima, J.P.; Krieger, I.E.; Krüger, F.P.; Lago, Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.A.; Authors: Martins, E.M.F.; Matsukawa, A.Y.; Mochizuki, C.F.M.; Minceda, E.C.; Miyake, C.Y.; P.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.A.; de Oliveira, F.P.; Pineda, F.A.; Rodrigues, V.; Rosa, A.J.; de M. de Rosa Jr., V.E.; de Sa, P.G.; Santilli, R.V.; Sawasak M.; Authors: da Silva, A.C.R.; da Silva, P.F.; da Silva, A.M.; Silva, J.; W.A.; da Silveira M.; Tsubako, M.B.; Valtada, B.; Van Sluys, M.A.; Verjovsky Almeida, S.; Vettore, A.L.; Z A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: Xf2242
C:Superfamily: regulatory protein spoJ

Query Match 0.6%; Score 61; DB 2; Length 264;
Best Local Similarity 27.9%; Pred. No. 69;
Matches 24; Conservative 12; Mismatches 50; Indels 0; Gaps 0;

QY 1595 TAVETIAPFKEENVLGEALCTAFVTSTFIAIAPQLPSMFSLTMTKNTSELVSSE 1754
Db 91 TAAATQMESESPFQPKPALSPISPIETETIILHPATSLILVNAITAAISVIVNOCF 150
QY 1755 VYLISALAAALOKVETLPHFISPLYE 1780
Db 151 YYALEGSSALLEETIEALKVNINORLE 176

RESULT 118
S40209
tubulin gamma chain - fungus (Cochliobolus heterostrophus)
C:Species: Cochliobolus heterostrophus, Bipolaris maydis
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 13-Aug-1999
C:Accession: S40209
R:Parkinson, C.; Luo, H.; Knight, A.; Ahlquist, J.; Perlin, M.B.
Submitted to the EMBL Data Library, August 1993
A:Description: Phylogenetic analyses using the gamma tubulin gene.
A:Reference number: S40209
A:Accession: S40209
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-265 <FAR>
A:Cross-references: EMBL:X74455; NID:q437988; PTDN:CAA52464.1; PID:q437989
C:Genetics:
A:Introns: 136;
C:Superfamily: tubulin

Query Match 0.6%; Score 61; DB 2; Length 265;
Best Local Similarity 28.6%; Pred. No. 70;
Matches 14; Conservative 11; Mismatches 24; Indels 0; Gaps 0;

QY 1407 GAGTVAGTFFDKVLTINMSIFTECAHVMRLDITYSFGVINKTVEMVTP 1355
Db 187 LVGIVASLITPTTGFALNTSTFSEWNEQAKTVKTVLVLMRFLQF 225

RESULT 119
A36704
bifunctional folate biosynthesis enzyme sub. Streptococcus pneumoniae
N:Alternate names: 6-hydroxymethyl-7,9-dihydropterin pyrophosphokinase, 7,9-dihydro-6-hydroxymethyl-2-amino-4-hydroxy-5-hydroxymethylidihydropteridine pyrophosphokinase (EC 2.7.1.1)
C:Species: Streptococcus pneumoniae
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A36704
R:Lopez, P.; Greenberg, B.; Lacks, S.A.
J. Bacteriol. 172, 4769-4774, 1990
A:Title: DNA sequence of folate biosynthesis gene sub., encoding hydroxymethylidihydropterin
A:Reference number: A36704; M01D:90368525
A:Accession: A36704
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-270 <LOP>

A:Cross-references: GR:M58706
C:Genetics:
A:Gene: sulD
C:Superfamily: bifunctional folate biosynthesis enzyme sub. 6-hydroxymethyl-7,9-dihydro-6-hydroxymethyl-2-amino-4-hydroxy-5-hydroxymethylidihydropteridine pyrophosphokinase (EC 2.7.1.1)
C:Species: Streptococcus pneumoniae
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A36704
R:Lopez, P.; Greenberg, B.; Lacks, S.A.
J. Bacteriol. 172, 4769-4774, 1990
A:Title: DNA sequence of folate biosynthesis gene sub., encoding hydroxymethylidihydropterin
A:Reference number: A36704; M01D:90368525
A:Accession: A36704
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-270 <LOP>

Query Match 0.6%; Score 61; DB 1; Length 270;
Best Local Similarity 35.3%; Pred. No. 71;
Matches 12; Conservative 6; Mismatches 16; Indels 0; Gaps 0;

QY 1747 TSELVSSEVYLLSALAAALOKVETLPHFISPLYE 1780
Db 225 TDSLILPHPIAPFLVLESLOELAPHFIIHPILK 258

RESULT 120
T21639
hypothetical protein F32B6.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C:Accession: T21639
R:Basham, V.
Submitted to the EMBL Data Library, October 1996
A:Reference number: Z19453
A:Accession: T21639
A:Status: preliminary; translated from GR/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-278 <WIL>
A:Cross-references: EMBL:Z81074; FIDN:CAH03283.1; GSPDB:GSD00022; GSPDB:F32B6.5
A:Experimental source: clone F32B6
C:Genetics:
A:Gene: CESP:F32B6.5
A:Map position: 4
A:Introns: 47/3; 204/1
C:Superfamily: Caenorhabditis elegans hypothetical protein F32B6.5

Query Match 0.6%; Score 61; DB 2; Length 278;
Best Local Similarity 22.4%; Pred. No. 74;
Matches 15; Conservative 13; Mismatches 39; Indels 0; Gaps 0;

QY 256 DYEAATYMLIQISVYVMTMTFFVNSLASQIIFITFTIPSLIFPHISQIVILQPKDPS 315
Db 23 EQPSEIVTVICIVIVEKAGPPTALNFKVSPVPTTAPVTAFFKKSSISPLAIGADQLIE 82

QY 316 LGPKDPP 322
Db 83 VTVSPVP 89

RESULT 121
T20726
hypothetical protein F10G8.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T20726
R:Basham, V.
Submitted to the EMBL Data Library, September 1996
A:Reference number: Z19315
A:Accession: T20726
A:Status: preliminary; translated from GR/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-296 <WIL>
A:Cross-references: EMBL:Z80216; FIDN:CAH02383.1; GSPDB:GSD00019; CESP:F10G8.7
A:Experimental source: clone F10G8
C:Genetics:
A:Gene: CESP:F10G8.7
A:Map position: 1
A:Introns: 53/1, 129/4, 199/4, 287/4, 353/1, 449/2

```

QY 1479 KIPFFKERTIPKAVSNKSEKSFPMLOVFNVTHTSKOLRHK 1514
|||||: 3 || | 3 || |
DB 212 KLPKQSSSEKFLVSNAPKSEKQKFFKFKKQLEDMCKKQKQIQ 254
|||||: 3 || | 3 || |

RESULT 135
T48463
hypothetical protein TIE3.30 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20 Apr 2000 #sequence_revision 20 Apr 2000
C:Accession: T48463
C:Author: N.J. Ardiles, W.J. Buyschaert, C.J. Passoville, R.J. De Clerck, R.J.
K.R. Bevan, M.J. Terry, S.J. Leach, K.J. Mayer, E.J.A.
ewes, H.W. Rudd, S.J. Leach, K.J. Mayer, E.J.A.
A:Reference number: Z24491
A:Accession: T48463
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1766 (REV)
A:Cross-references: EMBL:AL102972
A:Experimental source: cultivar Columbia; HAC clone TIE3
C:Genetics:
A:Map position: 5
A:Intens. 215.74; 452.74; 492.71
A:Note: TIE3.30

Query Match 0.68; Score 61; DH 2; Length 766;
Best Local Similarity 37.5%; Fred. NG. 2.40e-02;
Matches 12; Mismatch 6; Mismatch Pos 14; Indels 0; Gaps 0;

QY 814 NPFLKRSRDXLHLGLFEMMLNGADAVHF 845
| | | | | | | | | | | | | | | |
DB 571 NQKMLNEEDQVYLOQREAFELMIDGFKYHF 602
| | | | | | | | | | | | | | | |

```


C:Date: 30-Sep-1987 #sequence_revision 31 Mar 1991 #text_change 09 Jun 2000
 C:Accession: A23386; S06361
 R:Weber, J.L.; Leininger, W.M.; Lyon, J.A.
 Nucleic Acids Res. 14, 3311-3323, 1986
 A:Title: Variation in the gene encoding a major merozoite surface antigen of the human
 A:Reference number: A23386; MUID:86205236
 A:Accession: A23386
 A:Molecule type: DNA
 A:Residues: 1 1104 <WEB1>
 A:Cross-references: EMBL:X03831
 R:Weber, J.L.; Sim, B.K.L.; Lyon, J.A.; Wolff, R.
 Nucleic Acids Res. 16, 1706, 1988
 A:Title: Merozoite surface protein sequence from the Camp strain of the human malaria
 A:Reference number: S06361; MUID:88143999
 A:Accession: S06361
 A:Molecule type: DNA
 A:Residues: 1104-1726 <WEB2>
 A:Cross-references: EMBL:X03831
 C:Comment: The merozoite stages of different strains have strain-specific surface ant
 C:Comment: P. falciparum has three stages: sporozoite, merozoite, and gametocyte. The
 C:Superfamily: major merozoite surface antigen
 C:Keywords: glycoprotein; merozoite; surface antigen; tandem repeat
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-1726/protein: major merozoite surface antigen #status predicted <MA>
 F:67-87/91-96,105,109-120/Region: 3-residue repeats (S-G-T)
 F:757-765/Region: 3-residue repeats (T-E-E)
 F:133-272,591,597,607,827,833,934,936,1316,1334,1323,1613,1659/Binding site: carb

Query Match 0.68; Score 61; DB 2; Length 1726;
 Best Local Similarity 30.2%; Pred. No. 6,1e+02;
 Matches 16; Conservative 10; Mismatches 27; Indels 0; Gaps 0;

QY 958 SLAAVAVLQALFFELQSEKRSKQKLSLSEKLSVYSSSYAFLEMF 1020
 DB 1408 SSRAHDELSEKLSQKRTLLSKYKLEKVEDIPFAQGGISYVERVLAK 1460

RESULT 145
 A45948
 major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum)
 C:Species: Plasmodium falciparum
 C:Date: 30 Sep 1993 #sequence_revision 30-Sep-1993 #text_change 09 Jun 2000
 C:Accession: A45948
 R:Chang, S.P.; Kramer, K.J.; Yamaga, K.M.; Kato, A.; Case, S.E.; Siddiqui, W.A.
 Exp. Parasitol. 67, 1-11, 1988
 A:Title: Plasmodium falciparum, gene structure and hydropathy profile of the major me
 A:Reference number: A45948; MUID:89005525
 A:Accession: A45948
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1 1726 <CHIA>
 A:Cross-references: GB:M37213
 C:Superfamily: major merozoite surface antigen
 C:Keywords: surface antigen

Query Match 0.63; Score 61; DB 2; Length 1726;
 Best Local Similarity 30.2%; Pred. No. 6,1e+02;
 Matches 16; Conservative 10; Mismatches 27; Indels 0; Gaps 0;

QY 958 SLAAVAVLQALFFELQSEKRSKQKLSLSEKLSVYSSSYAFLEMF 1020
 DB 1408 SSRAHDELSEKLSQKRTLLSKYKLEKVEDIPFAQGGISYVERVLAK 1460

RESULT 146
 T29140
 hypothetical protein K11C4.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15 Oct 1999 #text_change 21 Jan 2000
 C:Accession: T29140
 R:Pauley, A.; Gattling, S.

Best Local Similarity 36.7%; Pred. No. 5,4e+02;
 Matches 11; Conservative 11; Mismatches 8; Indels 0; Gaps 0;

QY 1523 QLLSSQREKFKVVFQVPEPPEPPEPPEPPEPPEP 1552
 DB 563 RLINRRRLKRLKSSQAPOMIPENKPMQLQ 592

RESULT 142
 G70466
 DNA-directed RNA polymerase (EC 2.7.7.6) beta' chain - Aquifex aeolicus
 C:Species: Aquifex aeolicus
 C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 18-Jun-1999
 C:Accession: G70466
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, P.P.; O
 V.
 Nature 392, 353-358, 1998
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
 A:Reference number: A70300; MUID:98196666
 A:Accession: G70466
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1574 <AOE>
 A:Cross-references: GB:A600764, NID:42984196; FIDN:AA057724.1, PID:42984193, GR AP0006
 A:Experimental source: strain VF5
 A:Gene: rpoC
 C:Superfamily: Escherichia coli DNA-directed RNA polymerase beta' chain
 C:Keywords: nucleotidyltransferase; transcription

Query Match 0.68; Score 61; DB 2; Length 1574;
 Best Local Similarity 33.4%; Pred. No. 5,5e+02;
 Matches 11; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

QY 1520 FMSGLSSRNPKVWVSCQPHHKKLEPHELE 1552
 DB 399 FYRNVIRRRHLEKLEILNALGELLERKMLQ 431

RESULT 143
 A54498
 major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (is
 C:Species: Plasmodium falciparum
 C:Date: 28 Oct 1994 #sequence_revision 28-Oct-1994 #text_change 09 Jun 2000
 C:Accession: A54498
 R:Peterson, M.G.; Coppel, R.L.; McIntyre, P.; Langford, C.T.; Woodrow, G.; Brown, G.V.;
 Mol. Biochem. Parasitol. 27, 291-302, 1988
 A:Title: Variation in the precursor to the major merozoite surface antigens of Plasmodiu
 A:Reference number: A54498; MUID:88142999
 A:Accession: A54498
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1701 <PET>
 A:Cross-references: GB:M19143, NID:4160412; PID:AA29653.1; PID:G160413
 C:Superfamily: major merozoite surface antigen
 C:Keywords: surface antigen

Query Match 0.64; Score 61; DB 2; Length 1701;
 Best Local Similarity 30.2%; Pred. No. 6e+02;
 Matches 16; Conservative 10; Mismatches 27; Indels 0; Gaps 0;

QY 966 SLAAVAVLQALFFELQSEKRSKQKLSLSEKLSVYSSSYAFLEMF 1020
 DB 1383 SSRAHDELSEKLSQKRTLLSKYKLEKVEDIPFAQGGISYVERVLAK 1435

RESULT 144
 SA202M
 major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (st
 A:Alternate names: 195k glycoprotein
 C:Species: Plasmodium falciparum

A:Residues: 1-541 <ARN>
A:Cross-references: GB:AE061648, GB:AE061663, NID:q4376946, PID:AA018792.1; PID:q437695
A:Experimental source: strain CWI029
A:Genetic: R. K. Read, T. D. Brunkham, R. C. Shen, C. J. Gill, S. R. Heideberg, J. F. White, G. Hickey,
C. J. Jackson, R. G. Wilson, M. J. Nelson, M. J. DeBoy, K. J. Kennedy, C. J. McClarty, G. J. Saito, R. J.
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: AB1500; M01D:20150255
A:Accession: B81615
A:Molecule type: DNA
A:Residues: 1-541 <REA>
A:Cross-references: GB:AE062172, GB:AE062171, NID:J7189918, PID:AA037978.1; PID:J718992
A:Experimental source: strain AR39, HL cells
A:Genetic: C. Genetics:
A:Gene: cutE; CP0094
C:Superfamily: apolipoprotein N-acyltransferase

Query Match 0.6%; Score 60; DB 2; Length 541;
Best Local Similarity 33.3%; Pred. No. 2.2e+02;
Matches 13; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

QY 271 MRHFPVNSASQTEITETISLKEGSLIVLQPK 312
DB 74 LSDOYIGKLLVLMILITLSVLFSGSFIVAIVRQK 112

RESULT 168
B86572
apolipoprotein N-acyltransferase [imported] - Chlamydia pneumoniae (strain J138)
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C:Accession: B86572
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
Nucleic Acids Res. 30, 2211-2214, 2002
A:Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.
A:Reference number: AB6491; M01D:20340349
A:Accession: B86572
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-541 <STO>
A:Cross-references: GB:BA000008; NID:4897925, FIDR:BA09899.1; GSUP:GB060142
A:Experimental source: strain J138
A:Genetic: cutE

Query Match 0.6%; Score 60; DB 2; Length 541;
Best Local Similarity 33.3%; Pred. No. 2.2e+02;
Matches 13; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

QY 271 MRHFPVNSASQTEITETISLKEGSLIVLQPK 312
DB 74 LSDOYIGKLLVLMILITLSVLFSGSFIVAIVRQK 112

RESULT 169
T35245
probable ABC transporter - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 17-Mar-2000
C:Accession: T35245
R:Oliver, K.; Harris, E.; Bentley, S. D.; Parkhill, J.; Barrett, R. G.; Parkhill, J. M.
submitted to the EMBL Data Library, April 1999
A:Reference number: Z21573
A:Accession: T35245
A:Status: preliminary; translated from GR/EMBL/DBP1
A:Molecule type: DNA
A:Residues: 1-544 <OLI>
A:Cross-references: EMBL:AL049587; PID:CA04050.1; GSUP:GN00070, SC052A.03c
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SC052A.03c

C:Superfamily: unassigned ATP-binding cassette proteins: ATP-binding cassette homolog

Query Match 0.6%; Score 60; DB 2; Length 544;
Best Local Similarity 27.9%; Pred. No. 2.2e+02;
Matches 17; Conservative 10; Mismatches 34; Indels 0; Gaps 0;

QY 1040 QLEKIQKPTAVLKDEAMVHLTLGKYNEFSVLLNEPKSLDIFIKAVHTTKELYACM 1099
DB 272 QMYCTKPEPRVAVKLPATACQASACKYPIIMHFEKLAGAKERLDDAVEAVRDIETRVDL 331

QY 1100 P 1100
DB 332 P 332

RESULT 170
A45624
trophozoite cysteine proteinase - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jun-2000
C:Accession: A45624
P:Pasentbal, P. J.; Nelson, R. G.
Mol. Biochem. Parasitol. 51, 143-152, 1992
A:Title: Isolation and characterization of a cysteine proteinase gene of Plasmodium f
A:Reference number: A45624; M01D:92228005
A:Accession: A45624
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-565 <PO>
A:Cross-references: GB:M81341; GB:M85550, NID:J169247, PID:AAA9578.1; PID:q160248
A:Note: sequence extracted from RCB1 backbone (NCBI:98870, NCBI:98874)
C:Superfamily: trophozoite cysteine proteinase

Query Match 0.6%; Score 60; DB 2; Length 569;
Best Local Similarity 23.1%; Pred. No. 2.3e+02;
Matches 15; Conservative 20; Mismatches 30; Indels 0; Gaps 0;

QY 390 NLDHLIASLLFEYISYSSOEMDSKVSLLNQFLPLRLLESKYPTLDVLEPHLK 449
DB 162 MFTFEKYTLSETYNNKNNVSNIKNSIKSPFEYINLFFILERYKFFTFNNNEERK 161

QY 450 EIADL 454
DB 162 ELSNI 166

RESULT 171
D34541
phosphoprotein phosphatase (EC 3.1.3.16) 2A-beta 65K regulatory chain - pig (fragment
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 29-Jun-1990 #sequence_revision 29-Jun-1990 #text_change 24-Sep-1999
C:Accession: D34541
R:Hemmings, B. A.; Adams-Pearson, C.; Maurer, F.; Mueller, P.; Goris, J.; Morlevede, W
Biochemistry 29, 3165-3173, 1990
A:Title: Alpha- and beta-forms of the 65-kDa subunit of protein phosphatase 2A have a
A:Reference number: A34541; M01D:90241887
A:Accession: D34541
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-602 <HEM>
A:Cross-references: GB:734931, NID:q510346; PID:CAA84403.1; PID:q510347
C:Superfamily: phosphoprotein phosphatase 65K regulatory chain
C:Keywords: phosphoric monoester hydrolase

Query Match 0.6%; Score 60; DB 2; Length 602;
Best Local Similarity 22.2%; Pred. No. 2.4e+02;
Matches 16; Conservative 18; Mismatches 38; Indels 0; Gaps 0;

QY 2067 YQILLKTHISSPKVFAALLIVLALAEKLENYIVLLPESIFPLAEIMEDICEEVHQQC 2126
DB 1111

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure
A:Reference number: A75001
A:Accession: B75055
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-643 <RAW>
A:Cross-references: GB:AL096876; NID:q3458657; FID:CA852335.1; FID:GL51923
A:Experimental source: strain orsay
C:Genetics:
A:Gene: PAB1429

Query Match 0.6%; Score 60; DB 2; Length 643;
Best Local Similarity 37.9%; Pred. No. 2.7e+02;
Matches 11; Conservative 9; Mismatches 10; Indels 0; Gaps 0;
QY 729 RVPSILQKFKKLESVITAVEIPSEWHIE 756
DB 136 RVHSLQAKVQRIEALKRADIPFDLEIE 164

RESULT 177
G69177
Methionine--tRNA ligase (EC 6.1.1.10) - Methanobacterium thermoautotrophicum (strain Delta H)
A:Alternate names: methionyl-tRNA synthetase
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 16-Jul-1999
C:Accession: G69177
R:Smith, D.R.; Doucette Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Liu, D.; Spadafora, R.; Vicario, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.; K.; S.; Burch, C.M.; Burtis, C.J.; Mao, L.; Piro, P.; Neill, J.N.; Bacteriol. 179, 7135-7155, 1997
A>Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H, family
A:Reference number: Absam: M010:98037514
A:Accession: G69177
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-651 <MTB>
A:Cross-references: GB:AF060641; GB:AL060666; NID:q2521665; FID:AA85093.1; FID:gp22164
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH587

A:Start codon: TTG
A:Superfamily: methionine--tRNA ligase
C:Keywords: aminacyl tRNA synthetase; ligase; protein biosynthesis
Query Match 0.6%; Score 60; DB 2; Length 651;
Best Local Similarity 35.5%; Pred. No. 2.7e+02;
Matches 11; Conservative 9; Mismatches 11; Indels 0; Gaps 0;
QY 647 VIKSKIKKLEIGVAKRMIPDLAHTHIGDP 677
DB 611 VLVLKIKAKLEKVSFQVAVATPSNINLDP 641

RESULT 178
S63253
Probable membrane protein YNL279w - yeast (Saccharomyces cerevisiae)
A:Alternate names: hypothetical protein N0505
C:Species: Saccharomyces cerevisiae
C:Date: 27-Apr-1996 #sequence_revision 03-May-1996 #text_change 20-Jun-2000
C:Accession: S63253
R:Messey, F.; Dubois, E.; Vierendeels, F.; Scherens, H.; Pierard, A.; Glandsdorff, N.
A:Cross-references: EMBL:Z71555; FID:q1302348; FID:q1302349; GSPBR:GN05014; MIPS:YNL279w
A:Reference number: S63245
A:Accession: S63253
A:Molecule type: DNA
A:Residues: 1-661 <MES>
A:Cross-references: EMBL:Z71555; FID:q1302348; FID:q1302349; GSPBR:GN05014; MIPS:YNL279w
A:Experimental source: strain S288C
C:Genetics:
A:Gene: MIPS:YNL279w

A:Map position: 14L
C:Superfamily: Saccharomyces cerevisiae probable membrane protein YNL279w
C:Keywords: transmembrane protein
F:24-46/Domain: transmembrane #status predicted <TM>
F:110-126/Domain: transmembrane #status predicted <TM>
F:297-313/Domain: transmembrane #status predicted <TM>
F:422-448/Domain: transmembrane #status predicted <TM>
F:634-650/Domain: transmembrane #status predicted <TM>

Query Match 0.6%; Score 60; DB 2; Length 661;
Best Local Similarity 32.2%; Pred. No. 2.7e+02;
Matches 19; Conservative 10; Mismatches 30; Indels 0; Gaps 0;
QY 194 SLVTKSVKFAEYFGSSAVLKVLAFYASHIVSAVAALVSHIIRAKIRIVYKQKLES 252
DB 92 TLETLSLIVYASEGLVNFADLGLTYACLTIVSAVDGIVDAINTEKILISLVNDIVSS 150

RESULT 179
C71106
hypothetical protein PH0620 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 05-Nov-1999
C:Accession: C71106
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Hatakeyama, Y.; Hino, Y.; Yamamoto, S.; Se-
M.; Ohnaka, Y.; Yamahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kusuda, R.; Omi
DNA Res. 5, 55-76, 1998
A>Title: Complete sequence and gene organization of the genome of a hyperthermophilic
A:Reference number: A71000; M010:98344137
A:Accession: C71106
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-664 <RAW>
A:Cross-references: GB:AP060003; NID:q3236130; FID:BA22709.1; FID:dp06452; FID:p472
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenB
C:Genetics:
A:Gene: PH0620

Query Match 0.6%; Score 60; DB 2; Length 664;
Best Local Similarity 37.9%; Pred. No. 2.7e+02;
Matches 11; Conservative 8; Mismatches 10; Indels 0; Gaps 0;
QY 728 RVPSILQKFKKLESVITAVEIPSEWHIE 756
DB 156 RVHSLQAKVQRIEALKRADIPFDLEIE 184

RESULT 180
T36877
probable regulatory protein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T36877
R:Murphy, L.; Harris, F.; James, K.D.; Parkhill, J.; Farroll, R.C.; Palandream, M.A.
submitted to the EMBL Data Library, August 1999
A:Reference number: Z21617
A:Accession: T36877
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-682 <MBP>
A:Cross-references: EMBL:AL009448; FID:q4457836.1; GSHB:GN00070; SCODE:SC151.14C
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCODE:SC151.14C

Query Match 0.6%; Score 60; DB 2; Length 682;
Best Local Similarity 32.0%; Pred. No. 2.8e+02;
Matches 16; Conservative 9; Mismatches 25; Indels 0; Gaps 0;


```

      871 KCCCCCCCQGGHNCQUCHNINPGLALUCCCCCAGGAAV 916
      |...|...|...|...|...|...|...|...|...|...|
      872                                     RESULT 189

```

RESULT 189

[illegible]

Gibb's R.A., Myers E.W., Rubin G.M., Venter J.C.:
 "The genome sequence of *Escherichia coli* O157:H7".
 Science 287:2185-2195 (2000).
 -1- FUNCTION: REQUIRED FOR NUCLEAR PROTEIN IMPORT AND MEDIATES PACKING
 OF IMPORT SUBSTRATE TO DISTINCT NUCLEOPORINS.
 -1- SUBUNIT: FORMS A COMPLEX WITH AN IMPORTIN ALPHA SUBUNIT.
 -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 -1- SIMILARITY: BELONGS TO THE IMPORTIN BETA FAMILY.
 -1- SIMILARITY: CONTAINS 11 HEAT REPEATS.

 This SWISS-Prot entry is copyrighted. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation in
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement. For details, see <http://www.ebi.ac.uk/submit>
 or send an email to license@ebi.ac.uk.

 EMBL: A3092729; CAA05691.2; -
 EMBL: AF222745; AAF34680.1; -
 EMBL: AF003667; AAF53918.1; -
 FlyBase: FBpp000986; FS(2)Ket.
 InterPro: IPR000457; -
 InterPro: IPR001494; -
 PROSITE: PS50077; HEAT_REPEAT: 1.
 Transprot: Protein_Transport; Repeat.
 DOMAIN 129 489 HEAT REPEATS DOMAIN 1.
 DOMAIN 335 448 HEAT BINDING (BY SIMILARITY).
 DOMAIN 410 425 16X HEAT REPEATS (BY SIMILARITY).
 DOMAIN 606 742 HEAT REPEATS DOMAIN 2.
 SEQUENCE 884 AA; 98695 MW; D6DADCB927B414E2 PRO64;

 Query Match 3 73 Score 76; 15 in length 864;
 Best Local Similarity 35.68; Prod. No. 1.4;
 Matches 16; Conservative 11; Mismatches 10; Indels 0; Gaps 0;

 1522 SSSISNNRILKVVESQPELLEKPELLEIVLEKLSAVASQMF 1566
 22 RLKLSARILKLVASNNLELLKLSLLVNNLSAVAMNADIL 66

 RESULT 3
 VP13 YEAST
 ID VP13 YEAST STANDARD; PRT: 3144 AA.
 AC Q07876; Q00040;
 DT 01-NOV-1997 (Ref. 35, Created)
 DI 01-NOV-1997 (Ref. 45, Last sequence update)
 DI 01-OCT-2000 (Ref. 40, Last annotation update)
 DE VAPOR-LAB PROTEIN SORTING-ASSOCIATED PROTEIN VP13.
 GN VP13 OR SCL1 OR YL040C.
 OS *Saccharomyces cerevisiae* (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes;
 OX Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 GX NCBI_TaxID 4942;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN S288C;
 RA Brickner J.H., Fuller R.S.;
 RN Submitted (JUN 1997) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE OF 1-1460 FROM N.A.
 RA Duesterhoeft A., Floeth M., Heuss-Neitzel D., Hilbert H.,
 RA Moresl D.;
 RP Submitted (MAY 1996) to the EMBL/GenBank/DBJ databases.
 [3]
 RP SEQUENCE OF 1095-1144 FROM N.A.
 RA Wedler H., Wedler E., Scharle M., Wambull K.;
 RP Submitted (MAY 1997) to the EMBL/GenBank/DBJ databases.
 [4]
 RP CHARACTERIZATION.
 RA Luo W., Chang A.;

Fri Nov 9 15:42:41 2001

CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not altered. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/annotation/>

CC or send an email to license@sib-sib.ch).

EMBL; AB011676; BAA25136.1; -

EMBL; AB011674; BAA25134.1; -

PIR: A02955; DMPG.

InterPro; IPR001864; -

Pfam: PF00038; Filament; 1.

ProSITE; PS00226; IF; 1.

Intermediate filament; Coiled coil, heptad repeat pattern; Muscic.

INIT MET 0 0

DOMAIN 1 108

DOMAIN 139 412

DOMAIN 413 470

DOMAIN 109 141

DOMAIN 142 151

DOMAIN 152 252

DOMAIN 253 288

DOMAIN 269 287

DOMAIN 288 295

DOMAIN 296 412

SITE 354 354

DOMAIN 44 47

SEQUENCE 470 AA; 53497 MW; A2AEF7A3DBE5D512 CPG64;

Query Match 0.7%; Score 73; DB 1; Length 470;

Best local similarity 30.7%; Pred. No. 1-5;

Matches 23; Conservative 12; Mismatches 40; Indels 0; Gaps 0;

QY 2065 LRV...ILKTESEKFEFAALITVLALAKENYIVLPESIPFLAEIMDFQFVPHQ 2124

DB 194 LQEEIOLKEAEENLAFAKAVZAAATLARIIDLEEFESLDEIAPIKKVHPEFPIQAQ 253

QY 2125 CQTIQOETVLGEP 2139

DB 254 LQEQOVQVEMDMSKP 268

RESULT 8

EF3_PNECA STANDARD; PRT; 1042 AA.

AC P29551;

DT 01-APR-1993 (Rel. 25, Created)

DT 01-APR-1993 (Rel. 25, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE ELONGATION FACTOR 3 (EF-3).

GN TEF3.

OS Pneumocystis carinii.

OC Eukaryota; Fungi; Ascomycota, Pneumocystidomycetes; Pneumocystidaceae;

OC Pneumocystis.

OC NCBI_TaxID=4754;

RU [1]

RU SEQUENCE FROM N.A.

KX MEDLINE=93014112; PubMed=1339410;

RA Yipma Wang M., Ponzi W.A., Sypherd P.S.;

RT "Fungal-specific translation elongation factor 3 gene present in

PT Pneumocystis carinii";

CC Infert. Immun 60:4140-4145 (1992).

CC -1- FUNCTION: THE MAIN ROLE OF EF-3 MAY BE TO TRANSDUCE NUCLEOSIDE

CC TRIPHOSPHATE ENERGY INTO MECHANICAL ENERGY FOR TRANSLLOCATION

CC DURING TRANSLATION. EF-3 STIMULATES EF-1 ALPHA-DEPENDENT BINDING

CC OF AMINOACYL-TRNA TO THE RIBOSOME.

CC -1- PATHWAY: PROTEIN BIOSYNTHESIS.

CC -1- SUBUNIT: MONOMER.

CC -1- SIMILARITY: BELONGS TO THE ATT BINDING TRANSFER PROTEIN FAMILY

DOMAIN 108 411

DOMAIN 412 466

DOMAIN 108 140

DOMAIN 141 150

DOMAIN 151 251

DOMAIN 252 267

DOMAIN 268 286

DOMAIN 287 294

DOMAIN 295 411

VARIANT 172 178

VARIANT 336 336

VARIANT 344 344

VARIANT 359 359

VARIANT 392 392

CONFLICT 22 24

CONFLICT 38 48

CONFLICT 115 122

CONFLICT 134 134

SEQUENCE 469 AA; 53404 MW; 6A38116859A091B8 CPG64;

Query Match 0.7%; Score 73; DB 1; Length 469;

Best local similarity 30.7%; Pred. No. 1-5;

Matches 23; Conservative 12; Mismatches 40; Indels 0; Gaps 0;

QY 2065 LRV...ILKTESEKFEFAALITVLALAKENYIVLPESIPFLAEIMDFQFVPHQ 2124

DB 193 LQEEIOLKEAEENLAFAKAVZAAATLARIIDLEEFESLDEIAPIKKVHPEFPIQAQ 252

QY 2125 CQTIQOETVLGEP 2139

DB 253 LQEQOVQVEMDMSKP 267

RESULT 7

DESM_PIG STANDARD; PRT; 470 AA.

AC P02540; 062656;

DT 21-JUL-1986 (Rel. 01, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE DESMIN.

GN DES.

OS Sus scrofa (pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OC NCBI_TaxID=9823;

RU [1]

RU SEQUENCE FROM N.A.

KX TISSUE=Muscle;

RA Chikuni K., Tanabe K., Muroya S.;

RT "Desmin structure as related to meat tenderness";

RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.

RU [2]

RU SEQUENCE OF 333-470.

KX MEDLINE=82037793; PubMed=6945574;

RA Geisler N., Weber K.;

RT "Comparison of the proteins of two immunologically distinct

RT intermediate-sized filaments by amino acid sequence analysis: desmin

RT and vimentin";

RL proc. Natl. Acad. Sci. U.S.A. 78:4120-4123 (1981).

CC -1- FUNCTION: DESMIN ARE CLASS-III INTERMEDIATE FILAMENTS FOUND IN

CC MUSCLE CELLS. IN ADULT STRIATED MUSCLE THEY FORM A FIBROUS

CC NETWORK CONNECTING MYOFIBRILS TO EACH OTHER AND TO THE PLASMA

CC MEMBRANE FROM THE PERIPHERY OF THE Z-LINE STRUCTURES

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.

2

1- FUNCTION: THE P2A SUBUNIT OF PROTEIN PHOSPHATASE 2A SERVES AS A SCAFFOLDING MOLECULE TO COORDINATE THE ASSEMBLY OF THE CATALYTIC SUBUNIT AND A VARIABLE REGULATORY R SUBUNIT

1- SUBUNIT: P2A CONSISTS OF A COMMON HETERODIMERIC CORE ENZYME, COMPOSED OF A 36 KDA CATALYTIC SUBUNIT (SUBUNIT C) AND A 65 KDA CONSTANT REGULATORY SUBUNIT (P2A OR SUBUNIT A), THAT ASSOCIATES WITH A VARIETY OF DIMER INCLUSIVE THREE FAMILIES OF REGULATORY SUBUNITS

1- WITH THE CORE DIMER INCLUDE THREE FAMILIES OF REGULATORY SUBUNITS B (THE K2H/P255/455, P2-B, /P272/P216/P254 AND P5/B/456 FAMILIES), THE 4B KDA VARIABLE REGULATORY SUBUNIT, VIRAL PROTEINS, AND CELL SIGNALING MOLECULES.

1- DOMAIN: EACH HEAT REPEAT APPEARS TO CONSIST OF TWO ALPHA HELICES JOINED BY A HYDROPHILIC REGION, THE INTRAREPEAT LOOP, THE REPEAT UNITS MAY BE APPENDED LATERALLY TO FORM A 900-LIKE STRUCTURE

1- DISEASE: DEFECTS IN PPP2R1B COULD BE THE CAUSE OF SOME LUNG AND COLON CANCERS.

1- SIMILARITY: BELONGS TO THE PHOSPHATASE 2A REGULATORY SUBUNIT A FAMILY.

1- SIMILARITY: CONTAINS 15 HEAT REPEATS.

This SWISS PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/3cnouse/> or send an email to license@isb-sib.ch).

EMBL: AF083439; AAC63525.1; JOINED.
 EMBL: AF083425; AAC63525.1; JOINED.
 EMBL: AF083426; AAC63525.1; JOINED.
 EMBL: AF083427; AAC63525.1; JOINED.
 EMBL: AF083428; AAC63525.1; JOINED.
 EMBL: AF083429; AAC63525.1; JOINED.
 EMBL: AF083430; AAC63525.1; JOINED.
 EMBL: AF083431; AAC63525.1; JOINED.
 EMBL: AF083432; AAC63525.1; JOINED.
 EMBL: AF083433; AAC63525.1; JOINED.
 EMBL: AF083434; AAC63525.1; JOINED.
 EMBL: AF083435; AAC63525.1; JOINED.
 EMBL: AF083436; AAC63525.1; JOINED.
 EMBL: AF083437; AAC63525.1; JOINED.
 EMBL: AF083438; AAC63525.1; JOINED.
 EMBL: M65254; AAA59983.1; JOINED.
 EMBL: AF087458; AAC69624.1; JOINED.
 MIM: 603113.

PROSITE: PS50077; HEAT_REPEAT: 12.
 Multigene family; Repeat; Disease mutation.

HEAT 1.
 REPEAT 20 58
 REPEAT 59 96
 REPEAT 97 135
 REPEAT 136 173
 REPEAT 174 212
 REPEAT 213 251
 REPEAT 252 290
 REPEAT 291 333
 REPEAT 334 372
 REPEAT 373 411
 REPEAT 412 450
 REPEAT 451 489
 REPEAT 490 528
 REPEAT 529 567
 REPEAT 568 601
 REPEAT 602 640
 REPEAT 641 679
 REPEAT 680 718
 REPEAT 719 757
 REPEAT 758 796
 REPEAT 797 835
 REPEAT 836 874
 REPEAT 875 913
 REPEAT 914 952
 REPEAT 953 991
 REPEAT 992 1030
 REPEAT 1031 1069
 REPEAT 1070 1108
 REPEAT 1109 1147
 REPEAT 1148 1186
 REPEAT 1187 1225
 REPEAT 1226 1264
 REPEAT 1265 1303
 REPEAT 1304 1342
 REPEAT 1343 1381
 REPEAT 1382 1420
 REPEAT 1421 1459
 REPEAT 1460 1498
 REPEAT 1499 1537
 REPEAT 1538 1576
 REPEAT 1577 1615
 REPEAT 1616 1654
 REPEAT 1655 1693
 REPEAT 1694 1732
 REPEAT 1733 1771
 REPEAT 1772 1810
 REPEAT 1811 1849
 REPEAT 1850 1888
 REPEAT 1889 1927
 REPEAT 1928 1966
 REPEAT 1967 2005
 REPEAT 2006 2044
 REPEAT 2045 2083
 REPEAT 2084 2122
 REPEAT 2123 2161
 REPEAT 2162 2200
 REPEAT 2201 2239
 REPEAT 2240 2278
 REPEAT 2279 2317
 REPEAT 2318 2356
 REPEAT 2357 2395
 REPEAT 2396 2434
 REPEAT 2435 2473
 REPEAT 2474 2512
 REPEAT 2513 2551
 REPEAT 2552 2590
 REPEAT 2591 2629
 REPEAT 2630 2668
 REPEAT 2669 2707
 REPEAT 2708 2746
 REPEAT 2747 2785
 REPEAT 2786 2824
 REPEAT 2825 2863
 REPEAT 2864 2902
 REPEAT 2903 2941
 REPEAT 2942 2980
 REPEAT 2981 3019
 REPEAT 3020 3058
 REPEAT 3059 3097
 REPEAT 3098 3136
 REPEAT 3137 3175
 REPEAT 3176 3214
 REPEAT 3215 3253
 REPEAT 3254 3292
 REPEAT 3293 3331
 REPEAT 3332 3370
 REPEAT 3371 3409
 REPEAT 3410 3448
 REPEAT 3449 3487
 REPEAT 3488 3526
 REPEAT 3527 3565
 REPEAT 3566 3604
 REPEAT 3605 3643
 REPEAT 3644 3682
 REPEAT 3683 3721
 REPEAT 3722 3760
 REPEAT 3761 3799
 REPEAT 3800 3838
 REPEAT 3839 3877
 REPEAT 3878 3916
 REPEAT 3917 3955
 REPEAT 3956 3994
 REPEAT 3995 4033
 REPEAT 4034 4072
 REPEAT 4073 4111
 REPEAT 4112 4150
 REPEAT 4151 4189
 REPEAT 4190 4228
 REPEAT 4229 4267
 REPEAT 4268 4306
 REPEAT 4307 4345
 REPEAT 4346 4384
 REPEAT 4385 4423
 REPEAT 4424 4462
 REPEAT 4463 4501
 REPEAT 4502 4540
 REPEAT 4541 4579
 REPEAT 4580 4618
 REPEAT 4619 4657
 REPEAT 4658 4696
 REPEAT 4697 4734
 REPEAT 4735 4773
 REPEAT 4774 4812
 REPEAT 4813 4851
 REPEAT 4852 4890
 REPEAT 4891 4928
 REPEAT 4929 4967
 REPEAT 4968 5006
 REPEAT 5007 5045
 REPEAT 5046 5084
 REPEAT 5085 5123
 REPEAT 5124 5162
 REPEAT 5163 5201
 REPEAT 5202 5240
 REPEAT 5241 5279
 REPEAT 5280 5318
 REPEAT 5319 5357
 REPEAT 5358 5396
 REPEAT 5397 5435
 REPEAT 5436 5474
 REPEAT 5475 5513
 REPEAT 5514 5552
 REPEAT 5553 5591
 REPEAT 5592 5630
 REPEAT 5631 5669
 REPEAT 5670 5708
 REPEAT 5709 5747
 REPEAT 5748 5786
 REPEAT 5787 5825
 REPEAT 5826 5864
 REPEAT 5865 5903
 REPEAT 5904 5942
 REPEAT 5943 5981
 REPEAT 5982 6020
 REPEAT 6021 6059
 REPEAT 6060 6098
 REPEAT 6099 6137
 REPEAT 6138 6176
 REPEAT 6177 6215
 REPEAT 6216 6254
 REPEAT 6255 6293
 REPEAT 6294 6332
 REPEAT 6333 6371
 REPEAT 6372 6410
 REPEAT 6411 6449
 REPEAT 6450 6488
 REPEAT 6489 6527
 REPEAT 6528 6566
 REPEAT 6567 6605
 REPEAT 6606 6644
 REPEAT 6645 6683
 REPEAT 6684 6722
 REPEAT 6723 6761
 REPEAT 6762 6800
 REPEAT 6801 6839
 REPEAT 6840 6878
 REPEAT 6879 6917
 REPEAT 6918 6956
 REPEAT 6957 6995
 REPEAT 6996 7034
 REPEAT 7035 7073
 REPEAT 7074 7112
 REPEAT 7113 7151
 REPEAT 7152 7190
 REPEAT 7191 7229
 REPEAT 7230 7268
 REPEAT 7269 7307
 REPEAT 7308 7346
 REPEAT 7347 7385
 REPEAT 7386 7424
 REPEAT 7425 7463
 REPEAT 7464 7502
 REPEAT 7503 7541
 REPEAT 7542 7580
 REPEAT 7581 7619
 REPEAT 7620 7658
 REPEAT 7659 7697
 REPEAT 7698 7735
 REPEAT 7736 7774
 REPEAT 7775 7813
 REPEAT 7814 7852
 REPEAT 7853 7891
 REPEAT 7892 7930
 REPEAT 7931 7969
 REPEAT 7970 8008
 REPEAT 8009 8047
 REPEAT 8048 8086
 REPEAT 8087 8125
 REPEAT 8126 8164
 REPEAT 8165 8203
 REPEAT 8204 8242
 REPEAT 8243 8281
 REPEAT 8282 8320
 REPEAT 8321 8359
 REPEAT 8360 8398
 REPEAT 8399 8437
 REPEAT 8438 8476
 REPEAT 8477 8515
 REPEAT 8516 8554
 REPEAT 8555 8593
 REPEAT 8594 8632
 REPEAT 8633 8671
 REPEAT 8672 8710
 REPEAT 8711 8749
 REPEAT 8750 8788
 REPEAT 8789 8827
 REPEAT 8828 8866
 REPEAT 8867 8905
 REPEAT 8906 8944
 REPEAT 8945 8983
 REPEAT 8984 9022
 REPEAT 9023 9061
 REPEAT 9062 9100
 REPEAT 9101 9139
 REPEAT 9140 9178
 REPEAT 9179 9217
 REPEAT 9218 9256
 REPEAT 9257 9295
 REPEAT 9296 9334
 REPEAT 9335 9373
 REPEAT 9374 9412
 REPEAT 9413 9451
 REPEAT 9452 9490
 REPEAT 9491 9529
 REPEAT 9530 9568
 REPEAT 9569 9607
 REPEAT 9608 9646
 REPEAT 9647 9685
 REPEAT 9686 9724
 REPEAT 9725 9763
 REPEAT 9764 9802
 REPEAT 9803 9841
 REPEAT 9842 9880
 REPEAT 9881 9919
 REPEAT 9920 9958
 REPEAT 9959 9997
 REPEAT 10000 10038
 REPEAT 10039 10077
 REPEAT 10078 10116
 REPEAT 10117 10155
 REPEAT 10156 10194
 REPEAT 10195 10233
 REPEAT 10234 10272
 REPEAT 10273 10311
 REPEAT 10312 10350
 REPEAT 10351 10389
 REPEAT 10390 10428
 REPEAT 10429 10467
 REPEAT 10468 10506
 REPEAT 10507 10545
 REPEAT 10546 10584
 REPEAT 10585 10623
 REPEAT 10624 10662
 REPEAT 10663 10701
 REPEAT 10702 10740
 REPEAT 10741 10779
 REPEAT 10780 10818
 REPEAT 10819 10857
 REPEAT 10858 10896
 REPEAT 10897 10935
 REPEAT 10936 10974
 REPEAT 10975 11013
 REPEAT 11014 11052
 REPEAT 11053 11091
 REPEAT 11092 11130
 REPEAT 11131 11169
 REPEAT 11170 11208
 REPEAT 11209 11247
 REPEAT 11248 11286
 REPEAT 11287 11325
 REPEAT 11326 11364
 REPEAT 11365 11403
 REPEAT 11404 11442
 REPEAT 11443 11481
 REPEAT 11482 11520
 REPEAT 11521 11559
 REPEAT 11560 11598
 REPEAT 11599 11637
 REPEAT 11638 11676
 REPEAT 11677 11715
 REPEAT 11716 11754
 REPEAT 11755 11793
 REPEAT 11794 11831
 REPEAT 11832 11870
 REPEAT 11871 11909
 REPEAT 11910 11948
 REPEAT 11949 11987
 REPEAT 11988 12025
 REPEAT 12026 12064
 REPEAT 12065 12103
 REPEAT 12104 12142
 REPEAT 12143 12181
 REPEAT 12182 12220
 REPEAT 12221 12259
 REPEAT 12260 12298
 REPEAT 12299 12337
 REPEAT 12338 12376
 REPEAT 12377 12415
 REPEAT 12416 12454
 REPEAT 12455 12493
 REPEAT 12494 12532
 REPEAT 12533 12571
 REPEAT 12572 12610
 REPEAT 12611 12649
 REPEAT 12650 12688
 REPEAT 12689 12727
 REPEAT 12728 12766
 REPEAT 12767 12805
 REPEAT 12806 12844
 REPEAT 12845 12883
 REPEAT 12884 12922
 REPEAT 12923 12961
 REPEAT 12962 13000
 REPEAT 13001 13039
 REPEAT 13040 13078
 REPEAT 13079 13117
 REPEAT 13118 13156
 REPEAT 13157 13195
 REPEAT 13196 13234
 REPEAT 13235 13273
 REPEAT 13274 13312
 REPEAT 13313 13351
 REPEAT 13352 13390
 REPEAT 13391 13429
 REPEAT 13430 13468
 REPEAT 13469 13507
 REPEAT 13508 13546
 REPEAT 13547 13585
 REPEAT 13586 13624
 REPEAT 13625 13663
 REPEAT 13664 13701
 REPEAT 13702 13740
 REPEAT 13741 13779
 REPEAT 13780 13818
 REPEAT 13819 13857
 REPEAT 13858 13896
 REPEAT 13897 13935
 REPEAT 13936 13974
 REPEAT 13975 14012
 REPEAT 14013 14051
 REPEAT 14052 14090
 REPEAT 14091 14129
 REPEAT 14130 14168
 REPEAT 14169 14207
 REPEAT 14208 14246
 REPEAT 14247 14285
 REPEAT 14286 14324
 REPEAT 14325 14363
 REPEAT 14364 14401
 REPEAT 14402 14440
 REPEAT 14441 14479
 REPEAT 14480 14517
 REPEAT 14518 14556
 REPEAT 14557 14595
 REPEAT 14596 14634
 REPEAT 14635 14673
 REPEAT 14674 14711
 REPEAT 14712 14750
 REPEAT 14751 14789
 REPEAT 14790 14827
 REPEAT 14828 14866
 REPEAT 14867 14905
 REPEAT 14906 14944
 REPEAT 14945 14983
 REPEAT 14984 15022
 REPEAT 15023 15061
 REPEAT 15062 15100
 REPEAT 15101 15139
 REPEAT 15140 15178
 REPEAT 15179 15217
 REPEAT 15218 15256
 REPEAT 15257 15295
 REPEAT 15296 15334
 REPEAT 15335 15373
 REPEAT 15374 15412
 REPEAT 15413 15451
 REPEAT 15452 15490
 REPEAT 15491 15529
 REPEAT 15530 15568
 REPEAT 15569 15607
 REPEAT 15608 15646
 REPEAT 15647 15685
 REPEAT 15686 15724
 REPEAT 15725 15763
 REPEAT 15764 15802
 REPEAT 15803 15841
 REPEAT 15842 15879
 REPEAT 15880 15918
 REPEAT 15919 15957
 REPEAT 15958 15996
 REPEAT 15997 16035
 REPEAT 16036 16074
 REPEAT 16075 16113
 REPEAT 16114 16152
 REPEAT 16153 16191
 REPEAT 16192 16230
 REPEAT 16231 16269
 REPEAT 16270 16308
 REPEAT 16309 16347
 REPEAT 16348 16386
 REPEAT 16387 16425
 REPEAT 16426 16463
 REPEAT 16464 16502
 REPEAT 16503 16541
 REPEAT 16542 16580
 REPEAT 16581 16619
 REPEAT 16620 16658
 REPEAT 16659 16697
 REPEAT 16698 16735
 REPEAT 16736 16774
 REPEAT 16775 16813
 REPEAT 16814 16852
 REPEAT 16853 16891
 REPEAT 16892 16930
 REPEAT 16931 16969
 REPEAT 16970 17008
 REPEAT 17009 17047
 REPEAT 17048 17085
 REPEAT 17086 17123
 REPEAT 17124 17162
 REPEAT 17163 17200
 REPEAT 17201 17239
 REPEAT 17240 17278
 REPEAT 17279 17317
 REPEAT 17318 17355
 REPEAT 17356 17394
 REPEAT 17395 17433
 REPEAT 17434 17471
 REPEAT 17472 17510
 REPEAT 17511 17549
 REPEAT 17550 17588
 REPEAT 17589 17627
 REPEAT 17628 17665
 REPEAT 17666 17704
 REPEAT 17705 17743
 REPEAT 17744 17782
 REPEAT 17783 17821
 REPEAT 17822 17860
 REPEAT 17861 17899
 REPEAT 17899 17937
 REPEAT 17938 17976
 REPEAT 17977 18015
 REPEAT 18016 18054
 REPEAT 18055 18093
 REPEAT 18094 18132
 REPEAT 18133 18171
 REPEAT 18172 18210
 REPEAT 18211 18249
 REPEAT 18250 18288
 REPEAT 18289 18327
 REPEAT 18328 18366
 REPEAT 18367 18405
 REPEAT 18406 18444
 REPEAT 18445 18483
 REPEAT 18484 18522
 REPEAT 18523 18561
 REPEAT 18562 18600
 REPEAT 18601 18639
 REPEAT 18640 18678
 REPEAT 18679 18717
 REPEAT 18718 18755
 REPEAT 18756 18794
 REPEAT 18795 18833
 REPEAT 18834 18872
 REPEAT 18873 18910
 REPEAT 18911 18949
 REPEAT 18950 18988
 REPEAT 18989 19027
 REPEAT 19028 19066
 REPEAT 19067 19105
 REPEAT 19106 19144
 REPEAT 19145 19183
 REPEAT 19184 19222
 REPEAT 19223 19261
 REPEAT 19262 19300
 REPEAT 19301 19339
 REPEAT 19340 19378
 REPEAT 19379 19417
 REPEAT 19418 19456
 REPEAT 19457 19495
 REPEAT 19496 19534
 REPEAT 19535 19573
 REPEAT 19574 19612
 REPEAT 19613 19650
 REPEAT 19651 19689
 REPEAT 19690 19727
 REPEAT 19728 19766
 REPEAT 19767 19805
 REPEAT 19806 19844
 REPEAT 19845 19883
 REPEAT 19884 19922
 REPEAT 19923 19961
 REPEAT 19962 20000
 REPEAT 20001 20039
 REPEAT 20040 20078
 REPEAT 20079 20117
 REPEAT 20118 20156
 REPEAT 20157 20195
 REPEAT 20196 20234
 REPEAT 20235 20273
 REPEAT 20274 20312
 REPEAT 20313 20350
 REPEAT 20351 20389
 REPEAT 20390 20427
 REPEAT 20428 20466
 REPEAT 20467 20505
 REPEAT 20506 20544
 REPEAT 20545 20583
 REPEAT 20584 20621
 REPEAT 20622 20660
 REPEAT 20661 20699
 REPEAT 20699 20737
 REPEAT 20738 20776
 REPEAT 20777 20815
 REPEAT 20816 20854
 REPEAT 20855 20893
 REPEAT 20894 20931
 REPEAT 20932 20970
 REPEAT 20971 21009
 REPEAT 21010 21047
 REPEAT 21048 21086
 REPEAT 21087 21125
 REPEAT 21126 21163
 REPEAT 21164 21202
 REPEAT 21203 21241
 REPEAT 21242 21280
 REPEAT 21281 21319
 REPEAT 21320 21358
 REPEAT 21359 21397
 REPEAT 21398 21435
 REPEAT 21436 21474
 REPEAT 21475 21513
 REPEAT 21514 21552
 REPEAT 21553 21590
 REPEAT 21591 21629
 REPEAT 21630 21668
 REPEAT 21669 21706
 REPEAT 21707 21745
 REPEAT 21746 21783
 REPEAT 21784 21822
 REPEAT 21823 21860
 REPEAT 21861 21899
 REPEAT 21899 21937
 REPEAT 21938 21976
 REPEAT 21977 22015
 REPEAT 22016 22053
 REPEAT 22054 22092
 REPEAT 22093 22130
 REPEAT 22131 22169
 REPEAT 22170 22208
 REPEAT 22209 22247
 REPEAT 22248 22285
 REPEAT 22286 22323
 REPEAT 22324 22362
 REPEAT 22363 22400
 REPEAT 22401 22439
 REPEAT 22440 22478
 REPEAT 22479 22517
 REPEAT 22518 22555
 REPEAT 22556 22594
 REPEAT 22595 22632
 REPEAT 22633 22671
 REPEAT 22672 22709
 REPEAT 22710 22748
 REPEAT 22749 22787
 REPEAT 22788 22825
 REPEAT 22826 22864
 REPEAT 22865 22903
 REPEAT 22904 22941
 REPEAT 22942 22980
 REPEAT 22981 23019
 REPEAT 23020 23057
 REPEAT 23058 23096
 REPEAT 23097 23134
 REPEAT 23135 23173
 REPEAT 23174 23211
 REPEAT 23212 23250
 REPEAT 23251 23288
 REPEAT 23289 23327
 REPEAT 23328 23365
 REPEAT 23366 23404
 REPEAT 23405 23443
 REPEAT 23444 23481
 REPEAT 23482 23520
 REPEAT 23521 23559
 REPEAT 23560 23597
 REPEAT 23598 23636
 REPEAT 23637 23674
 REPEAT 23675 23713
 REPEAT 23714 23751
 REPEAT 23752 23790
 REPEAT 23791 23829
 REPEAT 23830 23868
 REPEAT 23869 23906
 REPEAT 23907 23945
 REPEAT 23946 23984
 REPEAT 23985 24022
 REPEAT 24023 24061
 REPEAT 24062 24099
 REPEAT 24100 24138
 REPEAT 24139 24177
 REPEAT 24178 24215
 REPEAT 24216 24254
 REPEAT 24255 24292
 REPEAT 24293 24331
 REPEAT 24332 24370
 REPEAT 24371 24409
 REPEAT 24410 24448
 REPEAT 24449 24486
 REPEAT 24487 24525
 REPEAT 24526 24563
 REPEAT 24564 24602
 REPEAT 24603 24640
 REPEAT 24641 24679
 REPEAT 24680 24717
 REPEAT 24718 24756
 REPEAT 24757 24794
 REPEAT 24795 24833
 REPEAT 24834 24871
 REPEAT 24872 24910
 REPEAT 24911 24948
 REPEAT 24949 24987
 REPEAT 24988 25025
 REPEAT 25026 25063
 REPEAT 25064 25102
 REPEAT 25103 25140
 REPEAT 25141 25179
 REPEAT 25180 25217
 REPEAT 25

H1704_ANAL11H

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See: <http://www.isb-sib.ch/amboss/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: M59443; AAA63542.1; -
 DR EMBL: AF1220; CA001244.1; -
 DR EMBL: M60460; AAA44103.1; -
 DR PIR: B38257; B38257.
 DR HSSP: P15737; IGHS.
 DR UniProt: IP0000490; -
 DR Pfam: PF00332; Glyco_hydro_17; 1.
 DR PROSITE: PS00587; GLYCOSYL_HYDROL_F17; 1.
 KW Hydrolase; Glycosidase, Signal; Multiene family.
 FT SIGNAL 1 29 PROBABLE.
 FT CHAIN 30 343 GLUCAN ENDO-1,3-BETA-GLUCOSIDASE, ACIDIC
 FT ISOFORM G19.
 FT MOD_RES 30 30 PYRROLIDONE CARBOXYLIC ACID (PROBABLE).
 FT ACT_SITE 264 264 NUCLEOPHILE (BY SIMILARITY).
 FT ACT_SITE 324 324 PROTON DONOR (BY SIMILARITY).
 FT CONFLICT 54 54 N > D (IN REF. 2).
 FT CONFLICT 172 172 Y -> N (IN REF. 2).
 FT CONFLICT 303 303 K -> N (IN REF. 2).
 FT CONFLICT 323 323 MISSING (IN AA SEQUENCE).
 FT SEQUENCE 343 AA; 37722 MW; 22854FFAF6B97E3 CRC64;
 SQ

Query Match 0.6%; Score 64; DB 1; Length 343;
 Best Local Similarity 31.0%; Pred No 15;
 Matches 13; Conservative 8; Mismatches 21; Indels 0; Gaps 0;

QY 1428 VLAAYGKDALEADTFWFSVCGFSGVQHQISLMNILOY 1469
 :||| 1111 3 3 3 1111 1111
 DB 99 ILANTPPKDSIFEGENSEINFLIQFLVQHNLFLANVVPY 140

RESULT 42
 E14G_TOBAC
 ID E136_TOBAC STANJAKG FKT 343 AA.
 AC P23547;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE GLUCAN ENDO-1,3-BETA-GLUCOSIDASE, ACIDIC ISOFORM G19 PRECURSOR
 DE (EC 3.2.1.39) ((1-3)-BETA-GLUCAN ENDOHYDROLASE) ((1-3)-BETA-
 DE GLUCANASE) (BETA-1,3-ENDOGLUCANASE) (PR-26) (PR-35).
 DE PR2.
 OS Nicotiana tabacum (Common Tobacco).
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliopsida; eudicotyledons; Core eudicotyledons; Asteridiales; Brassicidiales I.
 CC Solanales; Solanaceae; Nicotiana.
 CC NCBI_TaxID:4097;
 RN [1]
 RN SEQUENCE FROM N.A.
 RN STRAIN CV, SAMSON NN.
 RN MEDLINE 91062359, PubMed 2247145.
 RA Linthorst H.J.M., Melchers L.S., Mayer A., van Roekel J.S.C.,
 RA Cornelissen B.J.C., Bol J.F.;
 RA "Analysis of gene families encoding acidic and basic
 RA beta-1,3-glucanases of tobacco.";
 RA Proc. Natl. Acad. Sci. U.S.A. 87:8756-8760(1990).
 RA [2]
 RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RN STRAIN-CV, XANTINI NC; TISSUE-leaf,
 RA Ward P.P., Payne G.R., Meyer M.H., Williams S.C., Pincher S.S.,
 RA Sharkey F., Beck T.T., Taylor H.T., Abbot P., Melis P.,
 RA Ryals J.A.;
 RA "Differential regulation of beta-1,3-glucanase messenger RNAs in
 RA response to pathogen infection";
 RA Plant Physiol. 96:200-207(1991)
 RN [3]
 RN SEQUENCE OF 39-115, 157-158, 251-276 AND 328-343.
 RA van den Buijke M., Bauw G., Castresana G., van Montagu M.,
 RA Vanderkreeke J.,
 RA "Characterization of vascular and extracellular beta(1,3)-glucanases
 RA of tobacco: evidence for a strictly compartmentalized plant defense
 RA system.";
 RA Proc. Natl. Acad. Sci. U.S.A. 86:2674-2677(1989)
 CC 1-1- FUNCTION: IMPLICATED IN THE DEFENSE OF PLANTS AGAINST PATHOGENS.
 CC 1-1- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,3-BETA-D-GLUCOSIDIC LINKAGES
 CC IN 1,3-BETA-D-GLUCANS.
 CC 1-1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC 1-1- INDUCTION: NOT FOUND IN HEALTHY TISSUES, BUT ACCUMULATES TO HIGH
 CC LEVELS IN THE EXTRACELLULAR COMPARTMENT OF LEAVES IN RESPONSE TO
 CC PATHOGEN INFECTION OR TREATMENT WITH SALICYLIC ACID.
 CC 1-1- SIMILARITY: BELONGS TO FAMILY 17 OF GLYCOSYL HYDROLASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -


```

DR EMBL: K00922; AAA37104.1; JOINED.
DR EMBL: K00923; AAA37104.1; JOINED.
DR EMBL: K00924; AAA37104.1; JOINED.
DR EMBL: K00925; AAA37104.1; JOINED.
DR EMBL: K00926; AAA37104.1; JOINED.
DR EMBL: K00929; VHY.
DR HSSP: P03069; ISWI.
DR InterPro: IPR001684; -.
DR Pfam: PF00038; Filament; 1.
DR PROSITE: PS00226; IF; 1.
DR KW Intermediate filament; Coiled coil; Heptad repeat pattern;
KW Acetylation; Phosphorylation.
FT INIT_MET 0 0
FT DOMAIN 1 93 HEAD.
FT DOMAIN 94 405 ROD.
FT DOMAIN 406 464 TAIL.
FT DOMAIN 94 129 COIL 1A.
FT DOMAIN 130 151 LINKER 1.
FT DOMAIN 152 243 COIL 1B.
FT DOMAIN 244 266 LINKER 12.
FT DOMAIN 267 405 COIL 2.
FT MOD_RES 1 1 ACETYLATION.
FT MOD_RES 38 38 PHOSPHORYLATION (BY CAM-KINASE II)
FT MOD_RES 81 81 PHOSPHORYLATION (BY CAM-KINASE II)
FT MOD_RES 81 81 PHOSPHORYLATION (BY CAM-KINASE II)
FT CONFLICT 42 42 L -> A (IN REF. 2).
FT CONFLICT 115 115 Y -> D (IN REF. 2).
FT CONFLICT 182 182 R -> I (IN REF. 2).
FT SEQUENCE 464 AA; 53598 MW; 7AC417008CB04776 CRC64;

Query Match 0.64; Score 64; DB 1; Length 464;
Best Local Similarity 25.36; Pred. No. 22;
Matches 20; Conservative 18; Mismatches 41; Indels 6; Gaps 0;

QY 2061 LMKPIRFTLLRESETKVFAALIVLAIFPKRYVILNFSIPIFLMPPVPE 2129
DB 183 LREKLEMIQREAEATLQSFQDVQVNSLAPLNCPPVVSQREIAFLVPLHIDRIQF 242

QY 2121 VEHSCQFTQVQETVLGFF 2139
DB 243 LQAQIQPHVQIDVDVSKP 261

RESULT 36
VIME_HUMAN STANDARD: PRT; 465 AA.
AC P08670;
DI 01-JAN-1988 (Rel. 06, Created)
DI 01-OCT-1993 (Rel. 27, Last sequence update)
DI 01-NOV-1997 (Rel. 35, Last annotation update)
DE VIMENTIN.
GN VIM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxId=9606;
RN [1]
RP MEDLINE-87089701; PubMed-3467175;
RA Ferrari S., Battini R., Kaczmarek L., Pittling S., Calabretta R.,
RA de Riel J.K., Philipidis V., Wei J.F., Basvetta R.,
PT "Coding sequence and growth regulation of the human vimentin gene.",
RN Mol. Cell. Biol. 6:3614-3620(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-91067467; PubMed-2251132;
RA Honore B., Madsen P., Basse R., Andersen A., Walbum E., Celis J.E.,
RA Letters H.;
PT "Nucleotide sequence of cDNA covering the complete coding part of the
RN human vimentin gene.";
RN Nucleic Acids Res 18:6692-6692(1990).

```

```

[3]
RP SEQUENCE FROM N.A.
RA Zimbelmann R.;
FL Submitted (JAN-1994) to the EMBL Sequence DBPI databases.
[4]
RN SEQUENCE OF 16-24 AND 54-69.
KC TISSUE-Breast carcinoma;
RX MEDLINE-97295304; PubMed-9150946;
RA Pasmussen E.K., J.H., Pades J.S., Moritz R.L., Reid G.E.,
RA Simpson J., Dorow D.S.;
RT "Two-dimensional electrophoretic analysis of human breast carcinoma
PT proteins: mapping of proteins that bind to the SH3 domain of mixed
PT lineage kinase MLK2.";
PL Electrophoresis 18:588-598(1997).
CC -!- FUNCTION: VIMENTIN ARE CLASS-III INTERMEDIATE FILAMENTS FOUND IN
CC VARIOUS NONEPITHELIAL CELLS, ESPECIALLY MESENCHYMAL CELLS.
CC -!- SUBUNIT: HOMOPOLYMER.
CC -!- PTM: ONE OF THE MOST PROMINENT PHOSPHOPROTEINS IN VARIOUS CELLS OF
CC MESENCHYMAL ORIGIN. PHOSPHORYLATION IS ENHANCED DURING CELL
CC DIVISION, AT WHICH TIME VIMENTIN FILAMENTS ARE SIGNIFICANTLY
CC REORGANIZED.
CC -!- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.sib.ch.cam.ac.uk
CC or send an email to license@sib-sib.ch).
CC
DR EMBL: M14144; AAA61279.1; -.
DR EMBL: X56134; CAA39600.1; ALT_SEQ.
DR PMR: Z19554; CAA79613.1; -.
DR PIR: A25074; A25074.
DR PIR: S13115; S13115.
DR HSSP: P03064; ISWI.
DR Aarhus/Ghent-2DPAGE; 8417; IEF.
DR HSC-2DPAGE; P08670; HUMAN.
DR MIM: 193060; -.
DR Trp-P100; IPR001664; -.
DR Pfam: PF00038; filament; 1.
DR PROSITE: PS00226; IF; 1.
DR Intermediate filament; Coiled coil; Heptad repeat pattern.
KW Phosphorylation.
FT INIT_MET 0 0
FT DOMAIN 1 94 HEAD.
FT DOMAIN 95 406 ROD.
FT DOMAIN 407 465 TAIL.
FT DOMAIN 95 130 COIL 1A.
FT DOMAIN 131 152 LINKER 1.
FT DOMAIN 153 244 COIL 1B.
FT DOMAIN 245 267 LINKER 12.
FT DOMAIN 268 406 COIL 2.
FT MOD_RES 38 38 PHOSPHORYLATION (BY CAM-KINASE II)
FT MOD_PES R2 R2 (BY SIMILARITY).
FT CONFLICT 41 41 PHOSPHORYLATION (BY CAM-KINASE II)
FT CONFLICT 41 41 S -> D (IN REF. 1).
FT SEQUENCE 465 AA; 53554 MW; 858EC88732865FF1 CRC64;

Query Match 0.64; Score 64; DB 1; Length 465;
Best Local Similarity 32.04; Pred. No. 22;
Matches 16; Conservative 12; Mismatches 22; Indels 0; Gaps 0;

QY 2090 ALAEKLEKENVLLPESIPFLAKLEDECFHVPNGCKTLEQELVCEP 2139
DB 213 SLAPLDLPKVESIQFFIAFKLHEFEIQELQAIQEQHVQIDVDVSKP 262

RESULT 37
VIME_MOUSE

```



```

RX MEDLINE: 9217196; PubMed:1540169;
RA Bussemakers M.J.G., Verhaegh G.W.C.P., van der Vliet A.,
RA Debruyne F.M.J., Schaaijken J.A.;
RA "Differential expression of vimentin in rat prostatic tumours.";
RL Biochem. Biophys. Res. Commun. 182:1254-1259(1992).
RN 121
RP SEQUENCE OF 85 159 FROM N.A.
RA Patino M.L.;
RA Submitted (XXX-1992) to the EMBL/GenBank/DBPRT databases.
CC -!- FUNCTION: VIMENTIN ARE CLASS-III INTERMEDIATE FILAMENT'S FOUND IN
CC VARIOUS NONEPITHELIAL CELLS, ESPECIALLY MESENCHYMAL CELLS.
CC -!- SUBUNIT: HOMOPOLYMER.
CC -!- PTM: ONE OF THE MOST PROMINENT PHOSPHOPROTEINS IN VARIOUS CELLS OF
CC MESENCHYMAL ORIGIN. PHOSPHORYLATION IS ENHANCED DURING CELL
CC DIVISION, AT WHICH TIME VIMENTIN FILAMENTS ARE SIGNIFICANTLY
CC REORGANIZED.
CC -!- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC EMBL: X62952; CAA44722.1; -
CC EMBL: M84481, AAA42339.1; -
CC FIC, S22119, S22119.
CC FIC, JQ1389, JQ1389.
CC HSP: P03069; ISW1.
CC InterPro: IPR001664; -
CC Pfam: PF00038; filament; 1.
CC PROSITE: P00226; IF; 1.
KW Intermediate filament; Coiled coil; Heptad repeat pattern;
KW Phosphorylation.
FT INIT_MET 0 0 BY SIMILARITY.
FT DOMAIN 1 94 HEAD.
FT DOMAIN 95 406 ROD.
FT DOMAIN 407 465 TAIL.
FT DOMAIN 95 130 COIL 1A.
FT DOMAIN 131 152 LINKER 1.
FT DOMAIN 153 244 COIL 1B.
FT DOMAIN 245 267 LINKER 12.
FT DOMAIN 268 406 COIL 2.
FT MOD_RES 38 38 PHOSPHORYLATION (BY CAM-KINASE II)
FT FT 82 82 (BY SIMILARITY).
FT FT PROSGRGY:ATION (BY CAM-KINASE II)
FT FT (BY SIMILARITY).
FT SEQUENCE 465 AA; 53601 MW; 6587E7652CB682 CRC64;
Query Match 0.6%, Score 64; DB 1; Length 465;
Best Local Similarity 25.3%, Pred. No. 22;
Matches 20, Conservative 18; Mismatches 41; Indels 0; Gaps 0;
QY 2061 LWKPLRVGLIKTRESSTKRVNATLTVLAAELKERNVIVLLFSEIFLAELMEDECFE 2129
DB 184 LREKLQEMQREAEASTLQSFQVDNAGLNLRLDLKRVESLQETAPLKKCHDRETFQF 243
QY 2121 VRRGCKPTQQGLTVDLGF 2129
DB 244 LQAGVQGVQVDVSKP 262
QY 2121 VRRGCKPTQQGLTVDLGF 2129
DB 244 LQAGVQGVQVDVSKP 262
RESULT 39
DNAK_BACME STANISAKD, FET, 605 AA.
AC P05646;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE DNAK PROTEIN (HEAT SHOCK PROTEIN 70) (HSP70).

```

```

GN DNAK.
OS Bacillus megaterium
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
CC Bacillus/Staphylococcus group; Papillus
OX NCRI_TaxID=1404;
RN 111
RP SEQUENCE FROM N.A.
RX MEDLINE=87231083; PubMed=3035506;
RA Sussman M.D., Setlow P.;
RA "Nucleotide sequence of a Bacillus megaterium gene homologous to the
RA dnaK gene of Escherichia coli.";
RL Nucleic Acids Res. 15:3923-3923(1987).
CC -!- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).
CC -!- INDUCTION: BY HEAT SHOCK.
CC -!- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC EMBL: Y00154; CAA68348.1; -
CC HSP: P04475; 1DKX.
CC InterPro: IPR001023; -
CC Pfam: PF00012; HSP70; 1
CC PRINTS: PR00301; HEATSHOCK70.
CC PROSITE: PS00297; HSP70.1; 1
CC PROSITE: PS00329; HSP70.2; 1.
CC PROSITE: PS01036; HSP70.3; 1.
CC Chapterone; ATP-binding; Heat shock.
SQ SEQUENCE 605 AA; 65250 MW; 1CBCC2C4F34A9334 CRC64;
Query Match 0.6%, Score 64; DB 1; Length 605;
Best Local Similarity 32.7%, Pred. No. 31;
Matches 17, Conservative 12; Mismatches 23; Indels 0; Gaps 0;
QY 548 FSSEVTISNLNLFQRAELSKNGEWELVIAADILIKEITLSENDQLSNQV 599
DB 514 FTTEKTKIDLEGKVEAEVTKANFAKADKAAATFKNDLEETKAKKDELQELV 565
RESULT 40
Y228_BORBU STANISAKD, PPT, 471 AA
AC O51246;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE HYPOTHETICAL PROTEIN BR0228.
GN BR0228
OS Borrelia burgdorferi (lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCRI_TaxID=139;
RN 111
RP SEQUENCE FROM N.A.
RX MEDLINE=ATCC 35210 / B31;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey P.K., Swinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Karpavay A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt P., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Utterback T., Watthey L., McDonald L., Artach P., Howman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RA "Genomic sequence of a Lyme disease spirochaete, Borrelia
RA burgdorferi.";
PL Nature 390:580-586(1997).

```

CC THIS SWISS PROT ENTRY IS COPYRIGHTED. IT IS PRODUCED THROUGH A COLLABORATION
 CC BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION
 CC AT THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
 CC USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
 CC MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
 CC ENTITIES REQUIRES A LICENSE AGREEMENT (SEE <http://www.ebi.ac.uk/seqdb/doc/seqdb.html>)
 CC OR SEND AN EMAIL TO license@ebi.ac.uk.

DB EMBL: AF061150; AA:66621..1; -
 DR 1139; RB0228;
 KW Repeat local protein.
 SQ SEQUENCE 921 AA; 112959 MW; 988A688D73B6591A CR664;

Query Match 0.68; Score 64; DB 1; Length 971;
 Best local Similarity 54.18; Pred. No. 68;
 Matches 19; Conservative 9; Mismatches 18; Indels 0; Gaps 0;
 341 SEQFEALNLESLAKLEKSVKAVNKGQDENTSLFLHL 101
 879 SPASYRGNFKIVAFKSLKELANNMIDDELYLLGL 879

RESULT 41
 EMBL YEAST
 11 M84 YEAST STANDARD; PRT: 1089 AA.
 AC P54387;
 DT 01-07-1995 (Rel. 27, Created)
 DT 01-07-2000 (Rel. 43, Last sequence update)
 DT 01-07-2000 (Rel. 43, Last annotation update)
 DE IMPROVIN BETA 3 SUBUNIT (KAPY-RHEVIN BETA 3 SUBUNIT) (PE-PROTEIN
 DE SECRETION ENHANCER 1).
 OS PSEUDOMONAS AERUGINOSA (BACTERIA).
 ON Sacchariomyces cerevisiae (Baker's yeast).
 CC Eukaryotic; Group: Ascomycota; Sacchariomyces; Sacccharomycetes;
 CC Sacchariomyces; Sacccharomycetaceae; Sacccharomycetes;
 CC NCBI_TaxID:4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE 243641; PubMed:1522152;
 RA Chou L.Y.K., Ash J.J., Diamond D., Thomas D.Y.;
 RT "Screening and identification of a gene, PSE-1, that affects protein
 RT secretion in Saccharomyces cerevisiae.";
 RA J. Cell Sci. 101:709-719(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA SUPAIN 243641; PubMed:1522152;
 RA Samuel R., Dunbar J.M., Bartoll R.G., Rajandream M.A., Walsh S.V.;
 RT "Submitted (Apr 1999) to the EMBL/GenBank/DBJ databases."
 CC 1- FUNCTION: INVOLVED IN THE NUCLEAR IMPORT OF RIBOSOMAL PROTEINS.
 CC BELONGS TO NUCLEAR-IMPORTIN AND THE GTP-BIND FORM OF RSP1 (RAN).
 CC 1- FUNCTION: PLAYS A ROLE IN PROTEIN SECRETION.
 CC 1- SUBCELLULAR LOCATION: CYTOSOLIC AND NUCLEAR (BY SIMILARITY).
 CC 1- SIMILARITY: BELONGS TO THE IMPROVIN BETA FAMILY.

CC 1- SIMILARITY: CONTAINS 9 HEAT REPEATS.
 CC THIS SWISS PROT ENTRY IS COPYRIGHTED. IT IS PRODUCED THROUGH A COLLABORATION
 CC BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION
 CC AT THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
 CC USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
 CC MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
 CC ENTITIES REQUIRES A LICENSE AGREEMENT (SEE <http://www.ebi.ac.uk/seqdb/doc/seqdb.html>)
 CC OR SEND AN EMAIL TO license@ebi.ac.uk.

DB EMBL: Z11536; AA:77639..1; -
 DR 1139; RB0228;
 KW Transport; Protein transport; Nuclear protein.
 SQ SEQUENCE 921 AA; 112959 MW; 988A688D73B6591A CR664;
 341 SEQFEALNLESLAKLEKSVKAVNKGQDENTSLFLHL 101
 879 SPASYRGNFKIVAFKSLKELANNMIDDELYLLGL 879

FT REPEAT 404 441 HEAT 4;
 FT REPEAT 446 484 HEAT 5;
 FT REPEAT 489 526 HEAT 6;
 FT REPEAT 530 570 HEAT 7;
 FT REPEAT 719 756 HEAT 8;
 FT REPEAT 896 933 HEAT 9;
 FT CONFLICT 65 65 A -> S (IN REF. 1);
 FT CONFLICT 1077 1077 D -> A (IN REF. 1);
 SQ SEQUENCE 1089 AA; 121030 MW; 9C9F7F65B8B24C47 CR664;

Query Match 0.68; Score 64; DB 1; Length 1089;
 Best local Similarity 19.68; Pred. No. 68;
 Matches 19; Conservative 24; Mismatches 54; Indels 0; Gaps 0;
 341 SEQFEALNLESLAKLEKSVKAVNKGQDENTSLFLHL 101
 879 SPASYRGNFKIVAFKSLKELANNMIDDELYLLGL 879

RESULT 42
 EMBL RAT
 11 RB12 RAT STANDARD; PRT: 1145 AA.
 AC Q55081;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-07-2000 (Rel. 40, Last annotation update)
 DE RETINBLASTOMA-LIKE PROTEIN 2 (130 KDA RETINBLASTOMA ASSOCIATED
 DE PROTEIN) (PBL2) (P130) (RBR-2).
 OS Rattus norvegicus (rat).
 CC Eukaryotic; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurodonthi; Muridae; Rattus;
 CC NCBI_TaxID:10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN-SPRAGUE-DAWLEY; TISSUE:Spleen;
 RX MEDLINE=97384252; PubMed=9247086;
 RA Sawada Y., Nomura H., Endo Y., Umeki K., Fujita T., Ohaki S.,
 RA Fujinaga K.;
 RT "Cloning and characterization of the rat p130, a member of the
 RT retinoblastoma gene family.";
 RA Biochim. Biophys. Acta 1353:23-27(1997).
 CC 1- FUNCTION: MAY HAVE A FUNCTION IN CELL CYCLE REGULATION. BELONGS TO
 CC E1A PROTEIN. MAY ACT AS A TUMOR SUPPRESSOR. POTENT INHIBITOR OF
 CC E2F-MEDIATED TRANS-ACTIVATION. ASSOCIATES PREFERENTIALLY WITH
 CC E2F5. BINDS TO CYCLINS A AND E (BY SIMILARITY).
 CC 1- SUBCELLULAR LOCATION: NUCLEAR.
 CC 1- SIMILARITY: BELONGS TO THE RETINBLASTOMA PROTEIN (RB) FAMILY.

CC THIS SWISS PROT ENTRY IS COPYRIGHTED. IT IS PRODUCED THROUGH A COLLABORATION
 CC BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION
 CC AT THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
 CC USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
 CC MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
 CC ENTITIES REQUIRES A LICENSE AGREEMENT (SEE <http://www.ebi.ac.uk/seqdb/doc/seqdb.html>)
 CC OR SEND AN EMAIL TO license@ebi.ac.uk.

DB EMBL: D55627; AA:24196..1; -
 DR HSSP; P06400; LGDX.
 DR InterPro: IPR002719; -
 DR InterPro: IPR002720; -
 DR Pfam: PF01858; RB_A; 1;
 DR Pfam: PF01857; RB_B; 1;
 DR Transcription regulation; DNA-binding; Nuclear protein; Cell cycle;
 KW Phosphorylation; Anti-oncogene.
 FT DOMAIN 414 1021 BACKET (BINUS E1A).
 FT DOMAIN 414 613 DOMAIN A.

P49327;
01-FEB-1996 (rel. 33, Created)
01-FEB-1996 (rel. 33, Last sequence update)
01-OCT-2000 (rel. 40, Last annotation update)
FATTY ACID SYNTHASE (EC 2.3.1.85) [INCLUDES EC 2.3.1.38; EC 2.3.1.39;
EC 2.3.1.41; EC 1.1.1.103; EC 4.2.1.6]; E.C. 2.3.1.38; EC 2.3.1.39;
EC 2.3.1.41;
FASN OF FAS
HOMO sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxId=9606;
[1]
SEQUENCE FROM N.A..
TISUE=Brain;
MPBLIN=96064605; PubMed=7567999;
Gyakumar A., Tai M.-H., Huang W.-Y., Al-Feel W., Hsu M.,
Abu-Elheiga L., Chitala S.S., Wakil S.J.;
"Human fatty acid synthase: properties and molecular cloning";
Proc. Natl Acad Sci U S A 92:8695-8699(1995).
[2]
SEQUENCE OF 753-758 AND 1285-1297.
MEDLINE=94294385; PubMed 8622791;
Kubajda P., Jenner K., Wood F.D., Hennigar P.A., Jacobs L.B.,
Dick J.P., Pasternack G.R.;
"Fatty acid synthesis: a potential selective target for
antineoplastic therapy.";
Proc. Natl Acad Sci U S A 91:6379-6383(1994).
!- FUNCTION: FATTY ACID SYNTHASE CATALYZES THE FORMATION OF
LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.
THIS MULTIFUNCTIONAL PROTEIN HAS 7 CATALYTIC ACTIVITIES AND AN
ACYL CARRIER PROTEIN.
!- CATALYTIC ACTIVITY: ACETYL-COA + N MALONYL-COA + 2N NADPH ->
LONG CHAIN FATTY ACID + (N+1) COA + N CO(2) + 2H NADP(+).
!- CATALYTIC ACTIVITY: ACETYL-COA + [ACYL-CARRIER PROTEIN] = COA
+ ACETYL-[ACYL-CARRIER PROTEIN]
!- CATALYTIC ACTIVITY: MALONYL-COA + [ACYL-CARRIER PROTEIN] = COA
+ MALONYL-[ACYL-CARRIER PROTEIN].
! CATALYTIC ACTIVITY: ACYL-[ACYL-CARRIER PROTEIN] + MALONYL-[ACYL-
CARRIER PROTEIN] -> 3'-O-N-ACYL-[ACYL-CARRIER PROTEIN] + CO(2) +
[ACYL-CARRIER PROTEIN].
!- CATALYTIC ACTIVITY: (3R)-3-HYDROXYACYL-[ACYL-CARRIER PROTEIN] +
NADP(+) = 3'-O-N-ACYL-[ACYL-CARRIER PROTEIN] + NADPH
!- CATALYTIC ACTIVITY: (3R)-3-HYDROXYPALMITOYL [ACYL-CARRIER PROTEIN]
= 2-HEXADECENOYL-[ACYL-CARRIER PROTEIN] + H(2)O.
!- CATALYTIC ACTIVITY: ACYL-[ACYL-CARRIER PROTEIN] + NADP(+) -> 2-3-
DIHYDROXYACYL-[ACYL-CARRIER PROTEIN] + NADPH
!- CATALYTIC ACTIVITY: OLEOYL [ACYL-CARRIER PROTEIN] + H(2)O ->
ACYL-CARRIER PROTEIN + OLEATE.
!- SUBUNIT: HOMODIMER, WHICH IS REARRANGED IN A HEAD TO TAIL FASHION.
! TISSUE SPECIFICITY: PROMINENT EXPRESSION IN SPAIN, LUNG, AND
LIVER.
!- MISCELLANEOUS: THE RELATIVELY LOW BETA KETOACYL SYNTHASE ACTIVITY
MAY BE ATTRIBUTABLE TO THE LOW 4'-PHOSPHOPANTHETINE CONTENT OF
THE PROTEIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement ([see http://www.ebi.ac.uk/infocentre](http://www.ebi.ac.uk/infocentre)) or send an email to license@isb-sib.ch.

PMU; 026644; AAC0259.1; -
MIM; 600212; -
InterPro; IPR0000255; -
InterPro; IPR000794; -
InterPro; IPR01031; -
InterPro; IPR001227; -
InterPro; IPR002085; -
Pfam; PF00698; Acyl_transf; 1.
Pfam; PF00975; Thioesterase; 1.
Pfam; PF00107; adh_zinc; 1.

```

XX MEDLINE-91342636; PubMed-1875927;
XX Lock P., Ralph S., Stanley E., Boulet J., Ramsay P., Dunn A.P.;
XX "Two isoforms of marine hck, generated by utilization of alternative
XX translational initiation codons, exhibit different patterns of
XX subcellular localization.";
XX Mol. Cell. Biol. 11:4363-4370(1991).
XX [4]
XX X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 78-526.
XX MEDLINE-97177106; PubMed-9024658;
XX Sichert F., Moarefi I., Kuriyan J.;
XX "Crystal structure of the Src family tyrosine kinase Hck.";
XX Nature 385:602-609(1997)
XX [5]
XX STRUCTURE BY NMR OF 78-138.
XX MEDLINE-98239331; PubMed-9571048;
XX Horita D.A., Baldisseri D.M., Zhang W., Altieri A.S., Smithgall T.E.,
XX Gmeliner W.H., Byrd R.A.;
XX "Solution structure of the human Hck SH3 domain and identification of
XX its ligand binding site.";
XX J. Mol. Biol. 278:253-265(1998).
XX [6]
XX STRUCTURE BY NMR OF 139-245.
XX MEDLINE-97253487; PubMed-9109402;
XX Zhang W., Smithgall T.E., Gmeliner W.H.;
XX "Sequential assignment and secondary structure determination for the
XX Src homology 2 domain of hematopoietic cellular kinase.";
XX FEBS Lett. 406:141-145(1997).
XX [7]
XX X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 81-137.
XX MEDLINE-98453315; PubMed-9778343;
XX Arnold S., O'Brien P., Franken P., Strub M.P., Hoch P., Dumas C.,
XX Ladbury J.B.;
XX "RT loop flexibility enhances the specificity of Src family SH3
XX domains for HIV-1 Nef.";
XX Biochemistry 37:14583-14591(1998).
XX [8]
XX FUNCTION: MAY SERVE AS PART OF A SIGNALING PATHWAY COUPLING THE EC
XX RECEPTOR TO THE ACTIVATION OF THE RESPIRATORY BURST MAY ALSO
XX CONTRIBUTE TO NEUTROPHIL MIGRATION AND MAY REGULATE THE
XX DEGRANULATION PROCESS OF NEUTROPHILS.
XX [9]
XX CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
XX PROTEIN TYROSINE PHOSPHATE.
XX [10]
XX SUBCELLULAR LOCATION: P60-HCK AND P59-HCK ARE ASSOCIATED WITH
XX MEMBRANES. P60-HCK IS ALSO CYTOPLASMIC (BY SIMILARITY).
XX [11]
XX ALTERNATIVE PRODUCTS: THE P60-HCK AND P59-HCK ARE PRODUCED BY THE
XX USE OF ALTERNATIVE INITIATION SITES.
XX [12]
XX TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN CELLS OF THE
XX MYELOID AND B LYMPHOID LINEAGES.
XX [13]
XX SIMILARITY: CONTAINS 1 SH3 DOMAIN.
XX [14]
XX SIMILARITY: TO OTHER PROTEIN TYROSINE KINASES IN THE CATALYTIC
XX DOMAIN. BELONGS TO THE SRC SUBFAMILY.
XX
XX THIS SWISS PROT entry is copyright. It is produced through a collaboration
XX between the Swiss Institute of Bioinformatics and the EMBL outstation -
XX the European Bioinformatics Institute. There are no restrictions on its
XX use by non profit institutions as long as its content is in no way
XX modified and this statement is not removed. Usage by and for commercial
XX entities requires a license agreement (See http://www.isb-sib.ch/announce/
XX or send an email to license@isb-sib.ch).
XX
XX EMBL: M16591; AAA52643.1;
XX EMBL: M16592; AAA52644.1;
XX PIR: A27812; TVHUHC.
XX PDB: 2HCK; 2O-APG-97
XX PDB: 3HCK; 15-OCT-97.
XX PDB: 4HCK; 17-JUN-98
XX PDB: 5HCK; 17-JUN-98
XX PDB: 1A05; 15-MAY-97.
XX PDB: 1B01; 11-NOV-98.
XX MIN: 142370;
XX InterPro: IPR000719;
XX InterPro: IPR000980;
XX InterPro: IPR001245;

```

```

DR InterPro: IPR001452;
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF00018; SH3; 1.
DR Pfam: PF00069; PKINASE; 1.
DR PRINTS: PF00104; TYRKINASE.
DR PRINTS: PF00401; SH3DOMAIN.
DR PRINTS: PF00452; SH3DOMAIN.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS50001; SH2; 1.
DR PROSITE: PS50002; SH3; 1.
DR Trastelase; Tyrosine protein kinase; Phosphorylation; Atp-binding;
KW Myristate; SH2 domain; SH3 domain; Alternative initiation;
KW 3D-structure.
KW CHAIN 1 526 TYROSINE-PROTEIN KINASE P60-HCK.
FT CHAIN 22 526 TYROSINE-PROTEIN KINASE P59-HCK.
FT INIT_MET 22 22 POP P59-HCK.
FT LIPID 2 2 MYRISTATE (BY SIMILARITY).
FT LIPID 23 23 MYRISTATE (BY SIMILARITY).
FT DOMAIN 78 138 SH3.
FT DOMAIN 144 241 SH2.
FT DOMAIN 262 515 PROTEIN KINASE.
FT NP_BIND 288 276 ATP.
FT BINDING 290 290 ATP.
FT ACT_SITE 381 381
FT MOD_RES 411 411 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CONFLICT 24 24 S -> C (IN AAA52644).
SQ SEQUENCE 526 AA; 59583 MW; 347E877A0A6412B3 CRC64;

```

Query Match 0.68; Score 61; DB 1, Length 526;

Best Local Similarity 21.7%; Pred. No. 65;

Matches 16; Conservative 15; Mismatches 21; Indels 0; Gaps 0;

QY 1210 ILELQHKKRLRSQIIVPTLFLNLSNCTEPHPOGNGMVEYTKGLI 1255

DB 466 VIRALRGYRMPREPCPELYNIMMKCKNKKPKETFEYICSVL 511

RESULT 65

RYR2_SCHPO

ID RYR2_SCHPO STANDARD; PRT; 659 AA.

AC R28829;

DT 01-DEC-1992 (Rel. 24, Created)

DT 01-DEC-1992 (Rel. 24, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE PROTEIN KINASE RYR2 (EC 2.7.1.1) (PROTEIN KINASE SITE) (NAPE KINASE

DE KINASE) (MAPKK).

CN RYR2 OR STE8 OR SPRC107.05.

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota, Fungi, Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomyces.

OX NCBI_TaxID=4896;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE-91260705; PubMed-2846669;

RA Wang Y., Xu H.P., Riggs M., Rodgers L., Wiegler M.;

RA "pyr2, a Schizosaccharomyces pombe gene encoding a protein kinase

RA capable of partial suppression of the ras1 mutant phenotype.";

RT Mol. Cell. Biol. 11:3554-3563(1991).

RL [2]

RP SEQUENCE FROM N.A.

RA MEDLINE-9506795; PubMed-1435723.

RA Stykarsdottir U., Egel R., Nielsen O.;

RA "Functional conservation between Schizosaccharomyces pombe ste8 and

RA saccharomyces cerevisiae Ste11 protein kinases in yeast signal

RA transduction.";

RT Mol. Gen. Genet. 235:122-130(1992).

RL [3]

RP SEQUENCE FROM N.A.

RA [3]

RP STRAIN-972;


```

Coat protein; Glycoprotein.
W CHAIN          1   775      OUTER CAPSID PROTEIN VP4 .
T CHAIN          1   240      OUTER CAPSID PROTEIN VP6 .
T CHAIN          247   775      OUTER CAPSID PROTEIN VP5 .
T CARBOHYD       32    32        N-LINKED (GLCNAC) - ) (POTENTIAL) .
T CARBOHYD       56    56        N-LINKED (GLCNAC) - ) (POTENTIAL) .
T CARBOHYD       97    97        N-LINKED (GLCNAC) - ) (POTENTIAL) .
T CARBOHYD     129   129        N-LINKED (GLCNAC) - ) (POTENTIAL) .
T CARBOHYD     132   132        N-LINKED (GLCNAC) - ) (POTENTIAL) .
T CARBOHYD     195   195        N-LINKED (GLCNAC) - ) (POTENTIAL) .
T CARBOHYD     324   324        N-LINKED (GLCNAC) - ) (POTENTIAL) .
T CARBOHYD     583   583        N-LINKED (GLCNAC) - ) (POTENTIAL) .
T CARBOHYD     589   589        N-LINKED (GLCNAC) - ) (POTENTIAL) .
T CARBOHYD     592   592        N-LINKED (GLCNAC) - ) (POTENTIAL) .
T CARBOHYD     599   599        V -> I (IN REF. 2) .
T CONFLICT       35    35        MISSING (IN REF. 2) .
T CONFLICT      192   192        SL > PS (IN REF. 2) .
T SEQUENCE       236   237        GMA331C8A5D3PB3B CRC64 ;
SQ SEQUENCE       775 AA; 87687 MW; 6M431C8A5D3PB3B CRC64 ;

Query Match           0.6%, Score 61, DB 1, Length 775,
Best Local Similarity 30.2%; Pred. Re. 1e-02;
Matches 16; Conservative 11; Mismatches 26; Identities 0; Gaps 6;

ay 1045 LGKLTAVLKLEAMVLHLTLCKYRFEVELLHNEPKSLDFIKAVITTFPIVA 1047
| | | | | : | | | : | | | : | | | : | | | : | | | : | | | :
bb 667 LPEESLPFLFQVMEINTSEKFAKKITIMIVETGVRRFTTHLVNTTVISA 719

RESULT 68
IMBL_SCHPO STANDARD: PRT; 863 AA.
AC 013864;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 15-JUL-1998 (Rel. 36, last annotation update)
ID IMPORTIN BETA 1 SUBUNIT (KAJAYI HELEN BELLA SUBUNIT) (IMPORFIN 95).
GN SPAC1161.03c.
OS Schizosaccharomyces pombe (Fission yeast)
OC Eukaryota, Fungi, Ascomycota, Schizosaccharomycetes,
OC Schizosaccharomycetales, Schizosaccharomycetaceae;
OC schizosaccharomycetes
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Oliver K., Harris D., Barrell A.G., Rajandream M.A., Wood V.;
RE Submitted (Aug-1997) to the EMBL/GenBank/DDBB databases.
RL FUNCTION. REQUIRED FOR NUCLEAR PROTEIN IMPORT AND MEDIATES DOCKING
OF IMPORT SUBSTRATE TO DISTINCT NUCLEOPORINS. SERVES AS RECEPTOR
FOR NUCLEAR LOCALIZATION SIGNALS.
CC -1- SUBUNIT FORMS A COMPLEX WITH IMPORIN ALPHA SUBUNIT (BY
SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC AND NUCLEAR ENVELOPE (BY
SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE IMPORIN BETA FAMILY.
CC -1- SIMILARITY: CONTAINS 11 HEAT REPEATS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation at
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
```


Nature 392:353-358(1998).

-1- FUNCTION. DNA DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS SUBSTRATES.

-1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE + N PYROPHOSPHATE + RNA(N).

-1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1 BETA' CHAIN.

-1- SIMILARITY. BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb.sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: AF000764; AAC07724.1; .

InterPro: IPR000722; .

InterPro: IPR002474; .

PIfam: PF01854; RNA_pol_A2.1.

PIfam: PF00623; RNA_pol_A; 1.

Transfaser: DNA-directed RNA polymerase; Transcription.

SEQUENCE 1574 AA; 178500 MW; 4661AF291FEF7E78 CRC64;

Query Match 0.68; Score 61; DR 1; Length 1574;

Best Local Similarity 33.2%; Pred. NO. 2.4e+02;

Matches 11; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

Q7 1520 FMSCGLLEKREKESVSGEPELKEEFLEF 1562

Db 399 FYEVLNRENELEFLNAPRIILNREKMLQ 431

RESULT 72

MSPL-PLAF3

ID MSPL-PLAF3 STANDARD; PRT; 1682 AA.

AC Q19598; Q25921; .

DT 01-FEB-1991 (Rel. 17, Created)

DI 01-NOV-1997 (Rel. 35, Last annotation update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE MERZOITE SURFACE PROTEIN 1 (PEPUSOP) (MERZOITE SURFACE ANTIGENS) (PMMSA) (P190).

GN MSP-1.

OS Plasmodium falciparum (isolate ro-33 / Ghana).

OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.

OX NCBI_TaxID=5834; [1]

PP SEQUENCE OF 1-1061 FROM N.A.

RA MEDLINE=88166657; PubMed=3327688;

RA Ceita U., Kuhn D., Matic H., Reber Liske R.,

RT "A naturally occurring gene encoding the major surface antigen precursor p190 of Plasmodium falciparum lacks tripeptide repeats.";

RL EMBL J. 6:4137-4142(1987); [2]

RE SEQUENCE OF 1032-1682 FROM N.A.

RF MEDLINE=95354793; PubMed 7628569;

RA Toile R., Bujard H., Cooper J.A.;

RT "Plasmodium falciparum: variations within the C-terminal region of merozoite surface antigen-1.";

RL Exp. Parasitol. 81:47-54(1995);

CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR (POTENTIAL).

CC -1- PTM: MERZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 42 KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF MERZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: M35727; AAA29715.1; .

EMBL: Y00097; CAA68280.1; .

EMBL: Z35326; CAA84555.1; .

PIR: S06286; S06286.

InterPro: IPR000561; .

PIfam: PF00008; EGF; 1.

Malaria: Merzoite; Polypeptide; Signal; Glycoprotein;

KW Transmembrane; GPI-anchor.

FT SIGNAL 1 19

FT CHAIN 20 1682

FT TRANSMEM 166 1682

FT CARBOHYD 233 462

FT CARBOHYD 462 462

FT CARBOHYD 528 528

FT CARBOHYD 599 599

FT CARBOHYD 785 785

FT CARBOHYD 881 881

FT CARBOHYD 901 901

FT CARBOHYD 947 947

FT CARBOHYD 1071 1071

FT CARBOHYD 1178 1178

FT CARBOHYD 1569 1569

FT SEQUENCE 1682 AA; 194462 MW; 582A1E156943CAB6 CRC64;

Query Match 0.69; Score 61; DR 1; Length 1682;

Best Local Similarity 30.2%; Pred. NO. 2.7e+02;

Matches 16; Conservative 10; Mismatches 27; Indels 0; Gaps 0;

Q7 968 SDAAVYVQGLATFEELQPEKIKPSKIKSETFNITGVVVSQPSVIAFPI 1020

Db 1365 SSNEYITDSFKLNSEQKNTLLKSYKESVENDIKFAQEGISYKVLAK 1417

RESULT 73

MSPL-PLAFF

ID MSPL-PLAFF STANDARD; PRT; 1701 AA.

AC P13819; .

DT 01-JAN-1990 (Rel. 13, Created)

DI 01-JAN-1990 (Rel. 13, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE MERZOITE SURFACE PROTEIN 1 (PEPUSOP) (MERZOITE SURFACE ANTIGENS) (PMMSA).

GN MSP-1.

OS Plasmodium falciparum (isolate FC27 / Papua New Guinea).

OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.

OX NCBI_TaxID=5837; [1]

PP SEQUENCE FROM N.A

RA MEDLINE=88142994; PubMed=2449612;

RA Petersen M.G., Cooper R.L., McIntyre P., Langford C.J., Woodrow G.,

RA Brown G.V., Anders R.F., Kemp D.J.;

RT "Variation in the precursor to the major merozoite surface antigens of Plasmodium falciparum.";

RL M.L. Parasitol. 27:291-302(1988);

CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR (POTENTIAL).

CC -1- PTM: MERZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 42 KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF MERZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

[illegible]

T 15-JUL-1999 (Rel. 38, Last sequence update)
T 15-JUL-1999 (Rel. 38, Last annotation update)
E P/TUTITARY HOMEBOX 1 HOMOLOG (D-PHX1).
N PTX1.
S drosophila melanogaster (Fruit fly).
S Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
C Phryganea; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha,
C Ephydroidea; Drosophilidae; Drosophila.
X NCBI_TaxID=7227;
X [1]
N
P SEQUENCE FROM N.A.
C TISSUE=Embryo;
C MEDLINE=98092168; PubMed=9431811;
A Verbruggen G., Constien P., Zilian O., Wimmer F.A., Dowe G.,
A Taubert H., Noll M., Jaekle J.;
T "Embryonic expression and characterization of a pxl homolog in
T drosophila";
L Mech. Dev. 68:139-147(1997).
C [-] FUNCTION: MIGHT CONTROL PHYSIOLOGICAL CELL FUNCTIONS THAN PATTERNS
C FORMATION DURING EMBRYOGENESIS.
C [-] SUBCELLULAR LOCATION: Nuclear.
C [-] DEVELOPMENTAL STAGE: FIRST DETECTED IN THE POSTERIOR REGION OF
C THE BLASTODERM EMBRYO, IN LATER STAGES OF EMBRYONIC DEVELOPMENT,
C DETECTED IN THE POSTERIOR PORTION OF THE MIDGUT, IN THE
C DEVELOPING MALPHIGIAN TUBULES, IN A SUBSET OF VENTRAL SOMATIC
C MUSCLES, IN THE DEVELOPING CNS AND IN BOLWIG'S ORGAN.
C ! SIMILARITY: BELONGS TO THE PAIRED FAMILY OF HOMEBOX PROTEINS
C "HICOID" SUBFAMILY.

This SWISS-PROT entry is copyright © It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).

EMBL: AJ001519; CAA04801.1; -
Flybase: FBgc020912, ptx1.
InterPro: IPR001367.
Pfam: PF00046; homeobox_1.
PRINTS: PR00024; HOMEBOX.
PROSITE: PS00027; HOMEBOX_1; 1
PROSITE: PS00071; HOMEBOX_2; 1.
Homebox; DNA-binding; Developmental protein; Nuclear protein;
Transcription regulation; Activator.
DOMAIN 119 126 POLY-SER.
DNA_BIND 268 327 HOMEBOX.
SEQUENCE 513 AA; 54864 MW; 3EGB5C19CEB2E45 CRC64;

Query Match 0.6%; Score 60; DB 1; Length 513;
Best Local Similarity 30.6%; Prev No. 85;
Matches 11; Conservative 9; Mismatches 10; Indels 0; Gaps 0.

334 LHGISDYVSPLELYMLPHLVSIHHVTGETEG 369
+ + + | | : | | : | | : | | : |
116 LHDSSSSWSISPAISSIMPISLSHLHSAGQDLVG 151

RESULT 85
LNT_CHLPN LNT_CHLPN STANDARD; PRT; 541 AA.
AC Q9Z701; Q9JOG2;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
APOLIPROTEIN N-ACYLTRANSFERASE (EC 2.3.1.-) (ALP N-ACYLTRANSFERASE).
GN INT OR CUTR OR CPN0653 OR CP0094.
OS Chlamydia pneumoniae (Chlamydophilia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;

 This SWISS-PROT entry is copyright © 1993 and is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.sib.ch/announce/>) or send an email to license@isb.sib.ch.

EMBL: M35861; AAA34413.1; -
 EMRI: M31743; AAA35182.1; -
 EMBL: X16465; CAA4411.1; ALF_INIT.
 EMBL: X59720; CAA42259.1; -
 PIR: A36362; A36362.
 PIR: JN0133; JN0133.
 PIR: S11169; S11169.
 PIR: S19499; S19499.
 SGD: S0000680; TUPI.
 InterPro: IP0001680; -
 Pfam: PF00400; WD40; 7.
 PRINTS: PR00320; GPROTEINPROT.
 PROSITE: PS00678; WD_REPEATS_1; 4.
 PROSITE: PS50082; WD_REPEATS_2; 5.
 PROSITE: PS50294; WD_REPEATS_REGION; 1.
 Transcription regulation: Repeats: WD repeat.
 REPEAT 342 371 WD 1.
 REPEAT 441 471 WD 2.
 REPEAT 483 513 WD 3.
 REPEAT 524 555 WD 4.
 REPEAT 574 604 WD 5.
 REPEAT 628 658 WD 6.
 REPEAT 670 706 WD 7.
 DOMAIN 97 118 POLY-GIN.
 DOMAIN 181 198 THR-RICH.
 DOMAIN 399 409
 CONFLICT 75 75 E -> A (IN REF 1).
 CONFLICT 100 100 R -> Q (IN REF 1).
 CONFLICT 685 685 P -> S (IN REF. 1 AND 2).
 SEQUENCE 713 AA; 78307 MW; 444104AAD63CB944 CRC64;

Query Match 0.6%; Score 60; DB 1; Length 713;
 Best Local Similarity 35.6%; Fied. No. 1.3e+02;
 Matches 14; Conservative 10; Mismatches 16; Indels 0; Gaps 0;

Q7 1159 KQVPLEPPEKAKPLGTGTCQFPPEKMKQCKSKQNI ESQVE 1198
 DB 78 KHLKLT EOPHQTASITVCGQPCQCGQCGVGHLOQOQQ 117
 RESULT 94
 VP4_ROTND STANDARD; PRT; 775 AA.
 AC P11196;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)
 DE OUTER CAPSID PROTEIN VP4 (HEMAGGLUTININ) (OUTER LAYER PROTEIN VP4)
 DE [CONTAINS: OUTER CAPSID PROTEINS VP5 AND VP8].
 GN S4.
 OS Human rotavirus (serotype 2 / strain DSI).
 OC Viruses; dsRNA viruses; Reoviridae; Rotavirus
 OX NCBI_TaxID:10950;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE:88275070; PubMed:2849714;
 FX Garcia-Iglesias M., Green K., Nishikawa K., Taniguchi K., Jones R.,
 KA Kapikian A.Z., Chanock R.M.
 RT "Sequence of the fourth gene of human rotaviruses recovered from
 asymptomatic or symptomatic infections."
 RL J. Virol. 62:2978-2984 (1988).
 RN [2]
 RP SEQUENCE OF 1-280.

RA MEDLINE:86313706; PubMed:3018754;
 RA Garcia-Iglesias M., Hoshino Y., Becker White A., Rimentals I., Glass P.,
 Flores J., Kapikian A.Z., Chanock R.M.
 RT "Conservation of amino acid sequence of VP8 and cleavage region of
 VP4-VP8 outer capsid protein among rotaviruses recovered from
 asymptomatic neonatal infections."
 PL Proc. Natl. Acad. Sci. U.S.A. 83:7039-7043 (1986).
 CC -1- SUBCELLULAR LOCATION: OUTER CAPSID.
 CC -1- PTM: VP8 IS ONE OF TWO TRYPSIN CLEAVAGE PRODUCTS OF VP4; THE
 OTHER PRODUCT IS VP5.
 CC -1- SIMILARITY: BELONGS TO THE ROTAVIRUS VP4 PROTEIN FAMILY.

PIR: P28889; VPXPM4.
 PIR: C25904; VPXRDS.
 InterPro: IPR000416; -
 Pfam: PF00426; VP4; 1.
 Coat protein; Glycoprotein.
 CHAIN 1 775
 CHAIN 1 240
 FT CARBOHYD 56 56
 FT CARBOHYD 132 132
 FT CARBOHYD 150 150
 FT CARBOHYD 195 195
 FT CARBOHYD 324 324
 FT CARBOHYD 583 583
 FT CARBOHYD 589 589
 FT CARBOHYD 592 592
 FT CARBOHYD 599 599
 FT CONFLICT 52 53 HG -> SW (IN PFF 2).
 FT CONFLICT 106 107 IA -> SS (IN REF. 2).
 FT CONFLICT 142 144 MPK -> TPT (IN PFF 2).
 FT CONFLICT 245 245 P -> K (IN PFF 2).
 FT CONFLICT 280 280 V -> I (IN PFF 2).
 FT SEQUENCE 775 AA; 87592 MW; DFA5253DA451FE9 CRC64;

Query Match 0.6%; Score 60; DB 1; Length 775;
 Best Local Similarity 55.6%; Fied. No. 3.4e+02;
 Matches 16; Conservative 8; Mismatches 21; Indels 0; Gaps 0;

Q7 1045 IQKEPTAVLKDEAMVHLTLQKYNFVSLLNEDPKSLDIPKAV 1089
 DB 65/ 1PKFSYVVLKIGVEWMEINTECKFAKVKVDLTLEIEFDINKPAELV 711
 RESULT 94
 TGR3_PIG STANDARD; PRT; 848 AA.
 AC P35054;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE TGF-BETA RECEPTOR TYPE III PRECURSOR (TGFBR-3) (BETAGLYCAN).
 GN TGFBR3.
 OS Sus scrofa (pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID:9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE:Uterus;
 FX MEDLINE:94080582; PubMed:1333192;
 RA Moren A., Ichijo H., Miyazono K.
 RT "Molecular cloning and characterization of the human and porcine
 transforming growth factor-beta type III receptors."
 RL Biochem. Biophys. Res. Commun. 189:366-372 (1992).
 CC -1- FUNCTION: BINDS TO TGF-BETA. COULD BE INVOLVED IN CAPTURING AND
 RETAINING TGF-BETA FOR PRESENTATION TO THE SIGNALING RECEPTORS.
 CC -1- SURCELLULAR LOCATION: EXISTS BOTH AS A MEMBRANE-BOUND FORM AND AS
 SOLUBLE FORM IN SERUM AND IN THE EXTRACELLULAR MATRIX.
 CC -1- PTM: HEAVILY MODIFIED BY GLYCOSAMINGLYCAN GROUPS (GAS).
 CC -1- SIMILARITY: SOME REGIONS OF SIMILARITY TO ENDOGLIN.
 CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.

Query Match 0.5%; Score 59; DB 1; Length 243;
Best Local Similarity 29.2%; Pctd No. 47;
Matches 13; Conservative 13; Mismatches 21; Indels 0; Gaps 0;
QY 982 FELQPEKKLSKSHKSLSETIKNLISGVYSPSYIAKDIKMKVLCVNGTGM 1029
DB 33 KSMVEDSALTQNKISPTFGHLLAPQISPTDFPSIDLVFVIRKQMSL 180
RESULT 105
YA_ZYGFE
ID YA_ZYGFE STANDARD; PRT; 260 AA.
AC P13740;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-AUG-1990 (Rel. 15, Last annotation update)
DE HYPOTHETICAL PROTEIN A.
GN A.
OS Zygosaccharomyces fermentati
OC Plasmid pSM1.
GC Eukaryota, Fungi, Ascomycota; Saccharomycotina; Saccharomycetes;
CC Saccharomycetales, Saccharomycetaceae, Zygosaccharomycetes.
OX NCBI_TaxID=4955;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=NAGANISHI IPO 0021;
RX MEDLINE=88058763; PubMed=3060169;
RT Utatsu I., Sakamoto S., Imura T., Toh-E A.;
RT "Yeast plasmids resembling 2 micron DNA: regional similarities and
RT diversities at the molecular level.";
PL J. Bacteriol 169:5537-5545(1987).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.sib.ch/about-us>;
CC or send an email to license@isb.sib.ch).
CC -----
CC EMBL, M18275; AAA2276.1; -;
DE Hypothetical protein; Plasmid.
KW SEQUENCE 260 AA; 30438 MW; 0B68E5BA7073C13D CRC64;
Query Match 0.5%; Score 59; DB 1; Length 260;
Best Local Similarity 61.9%; Pctd No. 51;
Matches 13; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 979 TLFEELOPEKKLSKSHKSLSET 999
DB 204 TLFEELOPEKKLSKSHKSLSET 224
PFSUIT 106
NTRB_SALTY
ID NTRB_SALTY STANDARD; PRT; 349 AA.
AC P41788;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE NITROGEN PEPIDATION PROTEIN NR(1) (EC 2.7.3.-).
GN GNL OR NTRB.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC Salmonella.
OX NCBI_TaxID=602;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=LT2;
RX Kustu S.G.;
RA Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NTRB ACTS AS A SIGNAL TRANSDUCER WHICH RESPONDS TO THE

01-JUL-1993 (Rel. 26, Last sequence update)
10-MAY-2000 (Rel. 39, Last annotation update)
27 KDA MEMBRANE PROTEIN PRECURSOR (PEPTIDYL-PROLYL CIS-TRANS
ISOMERASE) (EC 5.2.1.8) (PIPIASE) (ROMAMASE A) (CHL-MIP)).
MIP OR CT541.
GN Chlamydia trachomatis.
OS Bacteria; Chlamydiales; Chlamydiales; Chlamydia;
NCBI_TaxID=813;
[1]
SEQUENCE FROM N.A.
STRAIN=SERVAR 12;
MEDLINE=94023883; PubMed=1406289;
Lundomose A.G., Rouch D.A., Birkelund S., Christiansen G.,
Pearce J.H.,
"Chlamydia trachomatis Mip-like protein";
Mol. Microbiol. 6:2539-2548(1992).
[2]
SEQUENCE FROM N.A.
STRAIN=DUN-57CX;
MEDLINE=95000809; PubMed=9784136.
Stephens R.S., Kalman S., Jammal C.J., Fan J., Marathe R., Aravind L.,
Mitchell W.P., Olinger U., Tarusov R.L., Zhao Q., Koonin E.V.,
Davis R.W.;
"Genome sequence of an obligate intracellular pathogen of humans:
Chlamydia trachomatis";
Science 292:754-759(1998)
[3]
SEQUENCE OF 69-243 FROM N.A.
STRAIN=SERVAR L2;
MEDLINE=91194539; PubMed=2013997;
Lundomose A.G., Birkelund S., Fey S.J., Mose Larsen F.,
Christiansen G.;
"Chlamydia trachomatis contains a protein similar to the Legionella
pneumophila mip gene product";
Mol. Microbiol. 5:109-115(1991)
-!- CATALYTIC ACTIVITY: CIS-TRANS ISOMERIZATION OF PROLINE IMIDIC
PEPTIDE BONDS IN OLIGOPPTIDES
-!- SUBCELLULAR LOCATION: OUTER MEMBRANE.
-!- SIMILARITY: BELONGS TO THE FKBP-TYPE PIPIASE FAMILY.

this SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb.sib.ch/about-us>;
or send an email to license@isb.sib.ch).

EMBL: X66126; CAA45917.1; -;
EMBL: X66127; CAA45918.1; -;
EMBL: X66128; CAA45919.1; -;
EMBL: AE001324; AAC68143.1; -;
PIR: S25255; S25255.
HSP: P20071; IPFK.
InterPro: IPR000774; -;
InterPro: IPR01179; -;
Pfam: PF00254; FKBP; 1.
Pfam: PF01346; FKBP_N; 1.
PROSITE: PS00453; FKBP_PP1ASE_1; 1.
PROSITE: PS00454; FKBP_PP1ASE_2; 1.
PROSITE: PS00059; FKBP_PP1ASE_3; 1.
outer membrane; Isomerase; Signal.
KW SIGNAL.
FT CHAIN 1 19
FT DOMAIN 20 243
FT DOMAIN 152 243
FT VARIAT 56 56
FT VARIAT 117 117
FT VARIAT 117 117
FT CONFLICT 56 56
FT CONFLICT 117 117
FT CONFLICT 137 138
FT CONFLICT 243 243
SEQUENCE 243 AA; 26548 MW; 2034F8C3F60BFD0C8 Cpc64;

NCBI_TaxID 46617;
 [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE-91218158; PubMed-1708831;
 RX Tian G.L., Michel F., Macadre C., Slonimski P.P., Lazowska J.;
 PT Principal mitochondrial evolution in yeasts II. The complete
 sequence of the gene coding for cytochrome b in Saccharomyces
 cerevisiae reveals the presence of both new and conserved introns and
 discloses major differences in the fixation of mutations in
 evolution.";
 RT J. Mol. Biol. 218:747-763(1991).
 RL [1]
 CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
 CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
 CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
 CC COUPLED TO ATP SYNTHESIS.
 CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
 CC BOUND TO THE PROTEIN.
 CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
 CC CYTOCHROME C1 AND THE RIESKE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
 CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X59280; CAA11971.1;
 CC InterPro: IPR000179;
 CC Pfam: PF00032; Cytochrome_b_2; 1.
 CC ProSite: PS00193; CYTOCHROME_B_HEME; 1.
 CC ProSite: PS00193; CYTOCHROME_B_OO; 1.
 CC Mitochondrion; Electron transport, Respiratory chain, Transmembrane;
 KW Heme.
 KW METAL 82 82 IRON 1 (HEME B562 AXIAL LIGAND).
 FT METAL 96 96 IRON 2 (HEME B566 AXIAL LIGAND).
 FT METAL 183 183 IRON 3 (HEME B562 AXIAL LIGAND).
 FT METAL 197 197 IRON 1 (HEME B566 AXIAL LIGAND).
 SQ SEQUENCE 385 AA; 43631 MW; 3541b2c7d77e2602 CRC64;
 Query Match 0.5%; Score 59; DB 1; Length 385;
 Best Local Similarity 45.3%; Pct. Id. 22;
 Matches 14; Conservative 5; Mismatches 12; Indels 0; Gaps 0;
 QY 1929 PRQELTFEYNLAQIAEKIKGILFLFAGHIV 1959
 ID 271 PEWLLDFYALIRSIPKLLGVITMFAAIIV 301
 RESULT 110
 CYR_YEAST
 ID CYR_YEAST STANDARD; PRT; 385 AA.
 AC 19563; Q95807; Q95801; Q55802;
 DT 21-JUL-1986 (Rel. 01, Created)
 DI 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE CYTOCHROME B,
 GN CYR_YEAST
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Mitochondrion.
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID 4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN D273-10B/A21;
 PC MEDLINE-81046788; PubMed-6254454;
 RA Reboreto F.G., Tzagoloff A.;
 RT "Assembly of the mitochondrial membrane system. DNA sequence and
 organization of the cytochrome b gene in Saccharomyces cerevisiae
 D273-10B.";
 RT J. Biol. Chem. 255:9828-9837(1980).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=D273-10B/A21;
 PC MEDLINE-85600973; PubMed-6381564;
 RA Bonjardim C.A., Nobrega F.G.;
 RT "Revision of the nucleotide sequence at the last intron of the
 RT mitochondrial apocytochrome b gene in Saccharomyces cerevisiae.";
 RT Braz. J. Med Biol. Res. 17:17-20(1984).
 RL [3]
 CC SEQUENCE FROM N.A.
 CC STRAIN=WR200;
 CC MEDLINE-95255283; PubMed-7737175;
 CC Claros M.G., Petrá J., Shu Y., Samatey F.A., Popot J.L., Jacq C.;
 CC "Limitations to in vivo import of hydrophobic proteins into yeast
 CC mitochondria. The case of a cytoplasmically synthesized apocytochrome
 CC b.";
 CC Eur. J. Biochem. 228:762-771(1995).
 RL [4]
 CC SEQUENCE OF 20-143 FROM N.A.
 CC STRAIN=777-3A;
 CC MEDLINE 81088336; PubMed 7001612;
 CC Lazowska J., Jacq C., Slonimski P.P.;
 CC "Sequence of introns and flanking exons in wild-type and box3 mutants
 CC of cytochrome b reveals an interlaced splicing protein coded by an
 CC intron.";
 CC Cell 22:333-348(1980).
 RL [5]
 CC SEQUENCE OF 144-160 FROM N.A.
 CC STRAIN=777-3A;
 CC MEDLINE-82113236; PubMed-7034963;
 CC Lazowska J., Jacq C., Slonimski P.P.;
 CC "Splice points of the third intron in the yeast mitochondrial
 CC cytochrome b gene.";
 CC Cell 27:12-14(1981).
 RL [6]
 CC MUTANT W7.
 CC MEDLINE-90005977; PubMed-2551731;
 CC Rivier-Chevillotte P., di Rago J.-P.;
 CC "Electron-transfer restoration by vitamin K3 in a complex
 CC III-deficient mutant of S. cerevisiae and sequence of the
 CC corresponding cytochrome b mutation.";
 CC FEBS Lett 255:5-9(1989).
 CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
 CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
 CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
 CC COUPLED TO ATP SYNTHESIS.
 CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
 CC BOUND TO THE PROTEIN.
 CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
 CC CYTOCHROME C1 AND THE RIESKE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY
 CC
 CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: J01476; AAA99924.1; ALT_SEQ.
 CC FMR1; X81042; CAA58863.1;
 CC FMR1; V00686; ; NOT_ANNOTATED_CDTS.
 CC EMBL: J01473; AAA32151.2;
 CC EMBL: J01472; AAA32151.2; JOINED
 CC EMBL: J01475; AAA32152.2;
 CC EMBL: J01474; AAA32152.2; JOINED
 CC PIR: A00159; CBBY.
 CC SCD: S0007270; CCB.
 CC InterPro: IPR000179; ;


```

DE 184 TERTIARY STRUCTURE OF THE NUCLEOCAPSID PROTEIN OF THE HUMAN PAPILLOMAVIRUS
QY 2121 VPEHGFQFTTQGLTWIGEP 2139
DB 244 LQAQTPQHVQIIMQVSKP 262

RESULT 113
ID VLI_HPV55 STANDARD: PRT: 501 AA.
AC P50420: G80940;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE MAJOR CAPSID PROTEIN 1
GN L1
OS Human papillomavirus type 55.
OC Virus; dsDNA viruses, no RNA stage; Papovaviridae; Papillomavirus
OX NCBI_TaxID: 57114;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H.;
RC Submitted (OCT-1995) to the EMBL/GenBank/DDBJ databases.
RN [2]
RN SEQUENCE OF 312-462 FROM N.A.
PX MEDLINE=95052821; PubMed=746466;
RA Bernard H.B., Chan S.Y., Marcos M.M., Ong C.K., Villa L.F.,
PA Delius H., Peyton C.L., Bauer H.M., Wheeler C.M.;
PT Identification and assessment of known and novel human
PT papillomaviruses by polymerase chain reaction amplification.
PT restriction fragment length polymorphisms, nucleotide sequence, and
PT phylogenetic algorithms.
RT J. Infect. Dis. 170:1077-1085(1994).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U31791; AAA79484.1;
CC EMBL: U12491; AAA67238.1;
CC InterPro: IPR002219;
CC Pfam: PF00500; late_protein_1;
CC PRINTS: PK00865; HPVcapsid1.
DR Coat protein; Late protein.
KW
SQ SEQUENCE 501 AA: 55787 MW: 0080407E552E968 CWC64;

Query Match 0.58; Score 59; DB 1; Length 501;
Best Local Similarity 25.48; Freq. No. 1.3e+02;
Matches 17; Conservative 11; Mismatches 39; Indels 0; Gaps 0;

QY 60 PSEGFQAPLSQAKLTPEVQTKVKNKGLDENISFLHLSPYFLKPKQKLEWLIH 119
DB 4 PSLAQVVIPLAFVSEVLTLLAVVEETHIVTHASSEFLAVGRIVYPAIFPAPKTLVPEVS 63

QY 120 KFHILY 126
DB 64 GFQYKVF 70

RESULT 114
NIKA_F0041 STANDARD: PRT: 524 AA.
AC NIK4_F0041
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE NICOTI RINGING PERIPLASMIC PROTEIN PRECURSOR
GN NIKA.

Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE=95020649; PubMed=7934931;
RA Navarro C., Wu L.-F., Mandrand-Berthelot M.-A.;
RT "The nik operon of Escherichia coli encodes a periplasmic binding-
RL protein-dependent transport system for nickel."
RN Mol. Microbiol. 9:1181-1191(1993).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=94316500; PubMed=8041620;
RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
RT "Analysis of the Escherichia coli genome. V. DNA sequence of the
RL Nucleic Acids Res. 22:2476-2588(1994).
CC -1- FUNCTION: INVOLVED IN A NICKEL TRANSPORT SYSTEM, PROBABLY
CC REPRESENTS THE NICKEL BINDER.
CC -1- SUBCELLULAR LOCATION: PERIPLASMIC (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING
CC PROTEIN FAMILY 5.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X73143; CAA51659.1;
CC EMBL: D00039; AAR18451.1;
CC EMBL: AF000423; AAC76650.1;
CC PIR: S3694; S36594.
CC ProGene: PGI2075; nika.
CC InterPro: IPR000014;
CC Pfam: PF00496; SBP_bac_5; 1.
CC PROSITE: PS01040; SBP_BACTERIAL_5; 1.
KW Transport; Nickel; Signal; Periplasmic.
FT SIGNAL 1 22
FT CHAIN 23 524 NICKEL-BINDING PERIPLASMIC PROTEIN.
SQ SEQUENCE 524 AA: 58719 MW: 06E3E3CE42396 7K764;

Query Match 0.58; Score 59; DB 1; Length 524;
Best Local Similarity 34.9%; Freq. No. 1.3e+02;
Matches 15; Conservative 7; Mismatches 21; Indels 0; Gaps 0;

QY 422 SLIVSALVAELVSNIIIAKIFPVIQKIKSSIFDYNAATYMI 264
DB 299 SLIDNALYGTQVATTFAPSPVPYANLGKESQYDPOKAKALL 341

RESULT 115
VHTJ_BPT7 STANDARD: PRT: 536 AA.
ID VHTJ_BPT7
AC P03728;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE HEAD-TO-TAIL JOINING PROTEIN.
GN 8.
OS Bacteriophage T7.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Podoviridae;
OC T7-like phages.
OX NCBI_TaxID=10760;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=84241725; PubMed=3864700;
RX

```


HAS2_HUMAN
Q92819;
30-MAY-2000 (rel. 39, Created)
30-MAY-2000 (rel 39, Last sequence update)
30-MAY-2000 (rel 39, Last annotation update)
HYALURONAN SYNTHASE 2 (EC 2.4.1.-) (HYALURONATE SYNTHASE 2)
(HYALURONIC ACID SYNTHASE 2) (HA SYNTHASE 2).
HAS2.
Homo sapiens (Human).
Enkaiyoga, Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi.

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homio.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
TISSUE=Brain;
MEDLINE=96394371; PubMed=8798477;
Watanabe K., Yamauchi Y.;
"Molecular identification of a putative human hyaluronan syn-
thetase." J. Biol. Chem. 271:22945-22948(1996).
-1- FUNCTION: PLAY A ROLE IN HYALURONAN/HYALURONIC ACID (HA)
SYNTHESIS.
-1- PATHWAY: HYALURONATE SYNTHESIS, WHICH HAS A REPEATING U
(CONSISTING OF GLUCURONIC ACID (GLUA) AND N-ACETYLGLUCALACT
(GLCNAC)).

This SWISS-PROT entry is copyright, it is produced through
between the Swiss Institute of Bioinformatics and the EMBL,
the European Bioinformatics Institute. There are no restrictions
use by non-profit institutions as long as its content
modified and this statement is not removed. Usage by and
entities requires a license agreement (See <http://www.ebi.ac.uk>
or send an email to license@ebi.sib.ch).

EMBL, D54804, AAC6062; -
MTM; 601636; -

Transferase: Glycosyl transferase	1	11	CYTOPLASMIC (POTENTIAL).
DOMAIN	12	32	1. (POTENTIAL).
TRANSFERM	33	45	EXTRACELLULAR (POTENTIAL).
DOMAIN	34	35	2. (POTENTIAL).
TRANSFERM	46	66	CYTOPLASMIC (POTENTIAL).
DOMAIN	67	374	3. (POTENTIAL).
TRANSFERM	375	395	EXTRACELLULAR (POTENTIAL).
DOMAIN	396	402	4. (POTENTIAL).
TRANSFERM	403	423	CYTOPLASMIC (POTENTIAL).
DOMAIN	424	439	5. (POTENTIAL).
TRANSFERM	440	450	EXTRACELLULAR (POTENTIAL).
DOMAIN	451	475	6. (POTENTIAL).
TRANSFERM	476	496	CYTOPLASMIC (POTENTIAL).
DOMAIN	497	510	7. (POTENTIAL).
TRANSFERM	511	531	EXTRACELLULAR (POTENTIAL).
DOMAIN	532	552	TRANSFERM
DOMAIN	552	AA:	EEFF5RD97R131FD CRC64;
SEQUENCE	552	AA:	53565 MW;
0.58;	Score 59;	DB 1;	Length 552;
Seq. Local Similarity:	28.64;	Prod. No. 1	30+02;
Catches 14;	Conservative	11;	Mismatches 24;
			Indels

```

973 VTODLATLFEELQREKKLKHQKLSSETLKNLLSCVSPSYAKDLMKV 102
    :|| : : : : : : : : : : : : : : : : : : : : : : :
58 IIGSLFAFLHRRMKKSLETPILKNKTVALCIAAQEHPDYLRKCSGV 106

```

P70312; P70411; Q62405;
 30-MAY-2000 (Rel. 39, Created)
 30-MAY-2000 (Rel. 39, Last sequence update)
 30-MAY-2000 (Rel. 39, Last annotation update)

1

DE HYALURONAN SYNTHASE 2 (EC 2.4.1.3) (HYALURONATE SYNTHASE 2)
 DE (HYALURONIC ACID SYNTHASE 2) (HA SYNTHASE 2)
 ON Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID:10090;
 PN SEQUENCE FROM N.A.
 RX STRAIN: C57BL/6J;
 RA Shaper A.P., Augustin M.L., McDonald J.A.;
 RT "Molecular cloning and characterization of a putative mouse hyaluronan
 synthase".
 RL J. Biol. Chem. 271:23400-23406(1996).
 FN 121
 RP SEQUENCE FROM N.A.
 RC STRAIN SWISS 311;
 RX MEGLINE 9716014; PubMed 879645;
 RA Futop C., Salustri A., Bascall V.C.;
 RT "Cloning sequence of a hyaluronan synthase homologue expressed during
 Arch. Biochem. Biophys. 337:261-266(1997).
 RN SEQUENCE 3,215-346 FROM N.A.
 RX STRAIN SWISS WERSER;
 RC MEGLINE 9720769; PubMed 864441;
 RA Semino S., Specht C.A., Palmieri A., Robbins P.W.;
 RT "Homologs of the Xenopus developmental gene Ect2 are present in
 cartilage and mouse and are involved in the development of cell line
 cell in osteoblasts during early embryogenesis".
 RL Proc. Natl. Acad. Sci. U.S.A. 94:4548-4553(1997).
 CN 1- FUNCTION: PLAY A ROLE IN HYALURONAN/HYALURONIC ACID (HA)
 SYNTHESIS.
 CC 1- PATHWAY: HYALURONATE SYNTHESIS, WHICH HAS A REPEATING UNIT
 CONSISTING OF GLUCURONIC ACID (GLCA) AND N-ACTYLGLACTOSAMINE
 (GLCANAC).
 CC 1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
 CC 1- SIMILARITY: BELONGS TO THE NUC/HAS FAMILY.
 CC This SWISS-prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 CC FM01: AF008201; AAR63209.1;
 PW Transferase; Glycosyltransferase; Transmembrane; Multigene family.
 FI DOMAIN 1 11 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 12 32
 FT DOMAIN 33 45 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 46 66
 FT DOMAIN 67 374
 FT TRANSMEM 375 395
 FT DOMAIN 396 402
 FT TRANSMEM 403 423
 FT DOMAIN 424 429
 FT TRANSMEM 430 450
 FT DOMAIN 451 475
 FT TRANSMEM 476 496
 FT DOMAIN 497 510
 FT TRANSMEM 511 531
 FT DOMAIN 532 552
 SQ SEQUENCE 552 AA; 63534 MW; 10CAA244E3A4E94C CRC64;

Query Match 0.5%; Score 59; DB 1; Length 552;
 Best Local Similarity 28.6%; Fred. No. 1-0002;
 Matches 14; Conservative 11; Mismatches 24; Indels 0; Gaps 0;
 Q7 47A VIQLALIFELHQLQKKKSLPTPIKTKVALCTAAVQEDPDYLRKIQSV 106
 DB 58 IIQSLFAFLHKKKKSLPTPIKTKVALCTAAVQEDPDYLRKIQSV 106

RESULT 120
 ID_HAS2_RAT
 AS HAS2_RAT
 AC 045776;
 DT 30-MAY-2000 (Ref. 39, last sequence update)
 DI 30-MAY-2000 (Ref. 39, last annotation update)
 DE HYALURONAN SYNTHASE 2 (EC 2.4.1.3) (HYALURONATE SYNTHASE 2)
 DE (HYALURONIC ACID SYNTHASE 2) (HA SYNTHASE 2).
 GN HAS2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN 11
 RP SEQUENCE FROM N.A.
 RA Sakaguchi K., Midura R.;
 RT "Molecular cloning of rat hyaluronan synthase 2 and regulation of its
 expression in a rat osteoblastic cell line".
 RL J. Biol. Chem. 271:1371-1377(1996).
 CN 1- FUNCTION: PLAY A ROLE IN HYALURONAN/HYALURONIC ACID (HA)
 SYNTHESIS (BY SIMILARITY).
 CC 1- PATHWAY: HYALURONATE SYNTHESIS, WHICH HAS A REPEATING UNIT
 CONSISTING OF GLUCURONIC ACID (GLCA) AND N-ACTYLGLACTOSAMINE
 (GLCANAC).
 CC 1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
 CC 1- SIMILARITY: BELONGS TO THE NUC/HAS FAMILY.
 CC This SWISS-prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 CC FM01: AF008201; AAR63209.1;
 PW Transferase; Glycosyltransferase; Transmembrane; Multigene family.
 FI DOMAIN 1 11 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 12 32
 FT DOMAIN 33 45 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 46 66
 FT DOMAIN 67 374
 FT TRANSMEM 375 395
 FT DOMAIN 396 402
 FT TRANSMEM 403 423
 FT DOMAIN 424 429
 FT TRANSMEM 430 450
 FT DOMAIN 451 475
 FT TRANSMEM 476 496
 FT DOMAIN 497 510
 FT TRANSMEM 511 531
 FT DOMAIN 532 552
 SQ SEQUENCE 552 AA; 63534 MW; 10CAA244E3A4E94C CRC64;

Query Match 0.5%; Score 59; DB 1; Length 552;
 Best Local Similarity 28.6%; Fred. No. 1-0002;
 Matches 14; Conservative 11; Mismatches 24; Indels 0; Gaps 0;
 Q7 47A VIQLALIFELHQLQKKKSLPTPIKTKVALCTAAVQEDPDYLRKIQSV 106
 DB 58 IIQSLFAFLHKKKKSLPTPIKTKVALCTAAVQEDPDYLRKIQSV 106

```

RA RA
RA Y256_HUMAN STANDARD; PRT; 635 AA.
AC Q93073;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HYDROTHERMAL PROTEIN KIAA0256.
GN KIAA0256.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RS SEQUENCE FROM N.A.
RC TISSUE=bone marrow.
RE MEDLINE=97191544; PubMed=9349592;
RA Nagase T., Seki N., Ishikawa K. I., Ohita M., Kawarabayashi Y.,
RA Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. VI.
RT The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
RT analysis of cDNA clones from cell line KC-1 and brain.";
RL DNA Res. 3:321-329(1996)
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@sib-sib.ch)
CC -----
EMBL: D87445; BAA13386.1;
DR Hypothetical protein.
KW SEQUENCE 635 AA; 70128 MW; 341820A024EF15E9 CPE54;
Query Match 0.5%; Score 59; DB 1; Length 635;
Best Local Similarity 29.5%, Prod. No. 1.5e+02;
Matches 13; Conservative 13; Mismatches 20; Indels 0; Gaps 0;
QY 953 YLIDILISRAEITSDAAVIGLIIILFELQREKKLSAQKL 956
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db 331 YFGAESLFNKLVELTEEARAYKUMVAAMEQEAEEALKNVKKV 374
RESULT 122
DPOL_SULSO
AC P26811; Q05707; STANDARD; PRT; 882 AA.
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE DNA POLYMERASE I (EC 2.7.7.7).
GN POLS.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RS SEQUENCE FROM N.A.
RC STRAIN=ATCC 49255 / MT4.
RE MEDLINE=92316966; PubMed=1614658.
RA Pisani F.M., Martino C., Rossi M.;
RT "A DNA polymerase from the archaeon Sulfolobus solfataricus shows
RT sequence similarity to family B DNA polymerases.";
RL Nucleic Acids Res. 20:2731-2736(1992)
RN [2]
RS SEQUENCE FROM N.A.
RC STRAIN=DSM 1617 / P2;
RE MEDLINE=97055432; PubMed=8899719;
RX Sensen C.W., Klenk H.-P., Singh R.K., Allard G., Chan C.C.-Y.,
PA

```

```

RA Liu Q.-Y., Penny S.L., Young F., Schenk M.E., Gaasterland T.,
RA Doolittle W.F., Ragan M.A., Charlebois R.L.;
RT "Organizational characteristics and information content of an
RT archaeal genome: 156 kb of sequence from Sulfolobus solfataricus
RT P2.";
RL Mol. Microbiol. 22:175-191(1996).
RN [3]
RS SEQUENCE FROM N.A.
RC STRAIN=DSM 1617 / P2;
RE MEDLINE=2016544R; PubMed=10701121;
RA Charlebois R.L., Singh R.K., Chan-Weiher C.C.-Y., Allard G., Chow C.,
RA Caftanlonieri T., Garrett P.A., Gordon P., Jeffries A.C., Kozera C.,
RA Gaasterland T., Garrett P.A., Gordon P., Penny S.L., She Q.,
RA Kushwaha N., Lafleur E., Medina N., Peng X., Penny S.L., She Q.,
RA St. Jean A., van der West J., Young F., Zivanovic Y., Doolittle W.F.,
RA Ragan M.A., Sensen C.W.;
RT "Gene content and organization of a 281-kbp contig from the genome of
RT the extremely thermophilic archaeon, Sulfolobus solfataricus P2.";
RL Genome. 43:116-136(2000)
CC 1 FUNCTION: THIS POLYMERASE POSSESSES TWO ENZYMAIC ACTIVITIES:
CC DNA SYNTHESIS (POLYMERASE) AND AN EXONUCLEOLYTIC ACTIVITY THAT
CC DEGRADES SINGLE STRANDED DNA IN THE 3' TO 5' DIRECTION.
CC 2 CATALYTIC ACTIVITY: NTP-->NMP+DIPHSIDE TRIPHOSPHATE.
CC 3 N PYROPHOSPHATE + DNA(N).
CC 4 SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@sib-sib.ch)
CC -----
EMBL: X64466; CAA45795.1;
DR EMBL: U92875; AAB53090.1;
DR EMBL: Y18930; CAA57747.1;
DR PIR: S23019; S23019
DR InterPro: IPR002064;
DR Pfam: PF00136; DNA_pol_B; 1.
DR PRINTS: PF00106; DNAPOLR
DR PROSITE: PS00116; DNA POLYMERASE B; 1.
KW Transferase; DNA-directed DNA polymerase; DNA replication;
KW RNA binding; Hydrolase; Exonuclease.
FT CONFLICT 139 139 L->H (IN REF. 1);
FT CONFLICT 624 624 A->R (IN REF. 1);
FT SEQUENCE 882 AA; 101223 MW; 77E3E575AA1C81A CPE64;
Query Match 0.5%; Score 59; DB 1; Length 882;
Best Local Similarity 30.8%, Prod. No. 2.2e+02;
Matches 16; Conservative 8; Mismatches 28; Indels 0; Gaps 0;
QY 67 APLFSQIAKTIERSVOTKAVNKOLDNISLFLHLSYFLLKPAQKCLEWLI 118
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db 622 APAVAESVIALGRYVITSTVEKAPFEGITVIVPTCSIFLNPENSLNLI 673
RESULT 123
SVL_BUCAI
AC P57249; STANDARD; PRT; 940 AA.
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ISO-EUCYL-TPNA SYNTHETASP (EC 6.1.1.5) (ISOLUCINE--TRNA LIGASE)
DE (HIFPS)
CN HIFPS OR B0149.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum)
OC symbiotic bacterium.
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=118099;
RN [1]

```


sclerosis patients.";
 Am. J. Hum. Genet. 62:286-294(1998).
 [12]
 RP VARIANTS TSC.
 RX MEDLINE=99045375; PubMed 9829910;
 RA Beauchamp R.L., Banwell A., McNamara P., Jacobsen M., Higgins E.,
 RA Northrup H., Short M.P., Sims K., Ozellus L., Ramesh V.,
 RA "Exon scanning of the entire TSC2 gene for germline mutations in 40
 RT unrelated patients with tuberous sclerosis.",
 PL Hum. Mutat. 12:408-416(1998).
 [13]
 RP VARIANTS TSC.
 RP Gilbert J.R., Guy V., Kumar A., Wolpert C., Kaudl R., Aylesworth A.,
 RA Roses A.D., Pericak-Vance M.A.;
 RA "Mutation and polymorphism analysis in the tuberous sclerosis 2 (TSC2)
 RT gene.",
 RL Neurogenetics 1:267-272(1999).
 [14]
 RP VARIANTS TSC, AND VARIANTS
 RX MEDLINE=9923423; PubMed 10205261;
 RA Jones A.C., Shyamundar M.M., Thomas M.W., Maynard J.,
 RA Griesbach S.A., Tomkins S., Sampson I.P., Cheadle J.P.;
 RA "Comprehensive mutation analysis of TSC1 and TSC2 and phenotypic
 RT correlations in 150 families with tuberous sclerosis.",
 RL Am. J. Hum. Genet. 64:1305-1315(1999).
 [15]
 RP VARIANT TSC ARG-717.
 RX MEDLINE=99167238; PubMed=10069705;
 RA Zhang H., Yamamoto T., Nanba E., Kikamura Y., Terada T., Akaboshi S.,
 RA Yuasa I., Ohtani K., Nakamoto S., Takeshita K., Ohno K.;
 RA "Novel TSC2 mutation in a patient with pulmonary tuberous sclerosis:
 RT lack of loss of heterozygosity in a lung cyst.",
 RL Am. J. Med. Genet. 82:368-370(1999).
 [16]
 RP VARIANTS TSC, AND VARIANTS
 RX MEDLINE=20197627; PubMed=10735580;
 RA Choy Y.S., Tabora S.T., Hall F., Ramesh V., Niida Y., Franz D.,
 RA Kasprzyk-Obara J., Peave M.P., Kwiatkowski D.J.;
 RA "Superiority of denaturing high performance liquid chromatography over
 RT single-stranded conformation and conformation sensitive gel
 RL electrophoresis for mutation detection in TSC2.",
 [17]
 RP VARIANTS TSC, AND VARIANTS
 RX MEDLINE=20067501; PubMed=10570913;
 RA Chang H., Nanba E., Yamamoto T., Ninomiya H., Ohno K., Mizuuchi M.,
 RA Takeshita K.;
 RA "Mutation analysis of TSC1 and TSC2 genes in Japanese patients with
 RT tuberous sclerosis complex.",
 RL J. Hum. Genet. 44:391-396(1999)
 [19]
 RP VARIANTS TSC W-611-W-905; P-1744 & H 1745 R 1751 DEL, & VARIANT L 320.
 RX TSC2=Periphereal blood leukocytes;
 RA MEDLINE=20076217; PubMed=10607950;
 RA Yamashita Y., Ono J., Okada S., Wataya-Kaneda M., Yoshikawa K.,
 RA Nishizawa M., Hirayama Y., Kobayashi E., Soyama K., Hino O.;
 RA "Analysis of all exons of TSC1 and TSC2 genes for germline mutations
 RT in Japanese patients with tuberous sclerosis: report of 10
 RL mutations.",
 RL Am. J. Med. Genet. 90:125-126(2000)
 CC -1- FUNCTION: IMPLICATED AS A TUMOR SUPPRESSOR. MAY HAVE A FUNCTION IN
 CC VESICULAR TRANSPORT, BUT MAY ALSO PLAY A ROLE IN THE REGULATION OF
 CC CELL GROWTH APPET AND IN THE REGULATION OF TRANSCRIPTION MEDIATED
 CC BY STEROID RECEPTORS. INTERACTION BETWEEN HAMARTIN AND TUBERIN MAY

FACILITATE VESICULAR DOCKING; SPECIFICALLY STIMULATES THE
 INTRINSIC GTPASE ACTIVITY OF THE RAS-RELATED PROTEIN RAP1A AND
 RAB5, SUGGESTING A POSSIBLE MECHANISM FOR ITS ROLE IN REGULATING
 CELLULAR GROWTH. MUTATIONS IN TUBERIN LEADS TO CONSTITUTIVE
 ACTIVATION OF RAP1A IN TUMORS.
 -1- SUBUNIT: INTERACTS WITH HAMARTIN. MAY ALSO INTERACT WITH THE
 ADAPTPP MOLECULE RAPPTIN 5. THE FINAL COMPLEX CONTAINS TUBERIN
 AND RAPPTIN 5 LINKED TO RAB5 (PROBABLE).
 -1- SUBCELLULAR LOCATION: CYTOPLASMIC. AT STEADY STATE FOUND IN
 ASSOCIATION WITH MEMBRANES.
 -1- ALTERNATIVE PRODUCTS: A NUMBER OF FORMS CAN BE DERIVED FROM
 ALTERNATIVE SPLICING OF THE TSC2 GENE.
 -1- TISSUE SPECIFICITY: LIVER, BRAIN, HEART, LYMPHOCYTES, FIBROBLASTS,
 BILIARY EPITHELIUM, PANCREAS, SKELETAL MUSCLE, KIDNEY, LUNG AND
 PLACENTA.
 -1- DISEASE: DEFECTS IN TSC2 ARE THE CAUSE OF TUBEROUS SCLEROSIS
 COMPLEX (TSC). THE MOLECULAR BASIS OF WHICH IS A FUNCTIONAL
 IMPAIRMENT OF THE TUBERIN-HAMARTIN COMPLEX. TSC IS AN AUTOSOMAL
 DOMINANT MULTI-SYSTEM DISORDER THAT AFFECTS ESPECIALLY THE BRAIN,
 KIDNEYS, HEART, AND SKIN. TSC IS CHARACTERIZED BY HAMARTOMAS
 (BENIGN OVERGROWTHS PREDOMINANTLY OF A CELL OR TISSUE TYPE THAT
 OCCURS NORMALLY IN THE ORGAN) AND HAMARTIAS (DEVELOPMENTAL
 ANOMALIES OF TISSUE DIFFERENTIATION). CLINICAL SYMPTOMS CAN RANGE
 FROM BENIGN HYPOFICIENT MACULES OF THE SKIN TO PROFOUND MENTAL
 RETARDATION WITH INTRACTABLE SEIZURES TO PREMATURE DEATH FROM A

Query Match 0.58; Score 59; DB 1; Length 1807;
 Best Local Similarity 25.88; Pred. No. 5.3e+02;
 Matches 17; Conservative 17; Mismatches 32; Indels 0; Gaps 0;
 QY 123 IMS:FTFMGANNVPLQYTSQVNFVTKVMVLPALIGSDGDSIEVSFNVEIVVKITSV 1382
 DB 172 LSSEFLVLVLNVLKFNESCYLDFYAPMVQMICILCVRTASSVDIFVSIQVLDVAVCYNCL 231
 QY 1383 FVDALP 1388
 DB 232 PAESLP 237
 PFSUFT 129
 FRAP_HUMAN STANDARD: PRT: 2549 AA.
 AC P42345;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE FRAP-PAPAMYCIN ASSOCIATED PROTEIN (FRAP) (PAPAMYCIN TARGET PROTEIN).
 GN FRAP.
 OS Homo sapiens (Human)
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 PP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=9427729; PubMed=8008069;
 RA Brown E.J., Albers M.W., Shin T.B., Ichikawa K., Keith C.T.,
 RA Lane W.S., Schreiber S.L.;
 RT "A mammalian protein targeted by G1-arresting rapamycin-receptor
 RL complex.",
 RL Nature 369:756-758(1994)
 RN [2]
 PP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 2018-2112
 RX MEDLINE=96279639; PubMed=8662507;
 RA Choi J., Chen J., Schreiber S.L., Clardy J.;
 RT "Structure of the FRAP12-rapamycin complex interacting with the
 RL binding domain of human FRAP.",
 RL Science 273:239-242(1996).
 RN [3]
 PP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 2018-2112.
 RX MEDLINE=99190960; PubMed=10089303;
 RA Liang J., Choi J., Clardy J.;
 RT "Pellucid structure of the FRAP12-rapamycin-UKB ternary complex at 2.2

CC -1- FUNCTION: THE SMALL PROTEIN NS2A, NS2B, NS4A AND NS4B ARE
 CC HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE RELATED FUNCTION.
 CC NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
 CC -1- SUMMARY: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
 CC LIPID PROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
 CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
 CC PROTEIN C AND RNA.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M62321; AAA45676.1; --
 CC PIR: A39166; GNVWC3.
 CC RSP: P27958; LHEI.
 CC MEROPS: S29.001; --
 CC MEROPS: U39.001; --
 CC InterPro: IPR0000745; --
 CC InterPro: IPR001490; --
 CC InterPro: IPR002166; --
 CC InterPro: IPR002519; --
 CC InterPro: IPR002519; --
 CC InterPro: IPR002521; --
 CC InterPro: IPR002522; --
 CC InterPro: IPR002531; --
 CC InterPro: IPR002868; --
 CC Pfam: PF01560; HCV_NS1; 1.
 CC Pfam: PF01538; HCV_NS2; 1.
 CC Pfam: PF01006; HCV_NS4a; 1.
 CC Pfam: PF01506; HCV_NS4b; 1.
 CC Pfam: PF00998; HCV_RdRp; 1.
 CC Pfam: PF01543; HCV_Capsid; 1.
 CC Pfam: PF01542; HCV_core; 1.
 CC Pfam: PF01539; HCV_env; 1.
 CC InterPro: IPR002531; --
 CC Core protein; GAG protein; Envelope protein; Helicase; ATP-binding;
 CC Transmembrane; Nonstructural protein; Hydr. class; Serine protease
 CC INIT_MET 1 1
 CC
 CC CHAIN 1 115
 CC CHAIN 116 191
 CC CHAIN 192 283
 CC CHAIN 284 729
 CC CHAIN 730 1006
 CC CHAIN 1007 1615
 CC CHAIN 1616 1862
 CC CHAIN 1863 2913
 CC CHAIN 2914 3611
 CC CHAIN 3612 369
 CC TRANSMEM 347 369
 CC ACT_SITE 1083 1083
 CC ACT_SITE 1107 1107
 CC ACT_SITE 1165 1165
 CC NP_BIND 1230 1237
 CC SITE 1316 1319
 CC SITE 1316 1319
 CC CARBOHYD 196 196
 CC CARBOHYD 209 209
 CC CARBOHYD 234 234
 CC CARBOHYD 305 305
 CC CARBOHYD 417 417
 CC CARBOHYD 423 423
 CC CARBOHYD 430 430
 CC CARBOHYD 448 448
 CC CARBOHYD 476 476
 CC CARBOHYD 532 532
 CC CARBOHYD 540 540
 CC CARBOHYD 556 556
 CC CARBOHYD 576 576

FT CARBOHYD 624 624 N-LINKED (GLCNAC) (POTENTIAL)
 FT CARBOHYD 645 645 N-LINKED (GLCNAC) (POTENTIAL)
 FT CARBOHYD 2041 2041 N-LINKED (GLCNAC) (POTENTIAL)
 FT CARBOHYD 2077 2077 N-LINKED (GLCNAC) (POTENTIAL)
 FT CARBOHYD 2249 2249 N-LINKED (GLCNAC) (POTENTIAL)
 FT CARBOHYD 2364 2364 N-LINKED (GLCNAC) (POTENTIAL)
 FT CARBOHYD 2789 2789 N-LINKED (GLCNAC) (POTENTIAL)
 SQ SEQUENCE 3011 AA: 427117 MW: 45586447 Da P5AF9 CR664;
 Query Match 0.5%; Score 59; DB 1; Length 3011;
 Best Local Similarity 34.5%; Pred. No. 9; Gaps 0;
 Matches 10; Conservative 10; Indels 0; Gaps 0;
 QY 1821 VLDPAIKTKYKQTEKNWKNHMGPPMSILQ 1849
 DB 1745 VIAPAVTNWOKLETFWAKHWNFISGIQ 1773
 RESULT 132
 YHP9_YEAST STANDARD; PRT; 3744 AA.
 AC P36811;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE HYPOTHEICAL 433.2 KDA PROTEIN IN HXT5-NPK1 INTERGENIC REGION.
 GN YHR099W.
 OS Saccharomyces cerevisiae (Baker's yeast)
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPRAIN-S288C / AR972;
 PX MEDLINE=94378003; PubMed=8091229;
 PA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
 Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,
 Kucaba T., Hillier L., Jier M., Johnston L., Landston Y.,
 Latreille P., Louis E.J., Macri C., Mardis F., Meneses S., Mouser L.,
 Nhan M., Pitkin L., Riles L., St. Peter H., Trevaskis F., Vaughan K.,
 Vignati B., Wilcox L., Wohlman P., Waterston R., Wilson R.,
 Vaudin M.;
 RA "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
 XI."
 RL Science 265:2077-2082(1994).
 RT Science 265:2077-2082(1994).
 CC -1- SIMILARITY: STRONG, TO S.POMBE SPAC1F5.11C.
 CC -1- SIMILARITY: BELONGS TO THE PI3/P14-KINASES FAMILY.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U00060; AAB68923.1; --
 CC PIR: S46715; S46715.
 CC SGD: S0001141; YHR099W.
 CC KW Hypothetical protein; Transferase; Kinase.
 CC DOMAIN 3414 3744 P13K/PI4K.
 SQ SEQUENCE 3744 AA: 433171 MW: 45358676 Da P5P511 CR664;
 Query Match 0.5%; Score 59; DB 1; Length 3744;
 Best Local Similarity 50.0%; Pred. No. 1; Gaps 0;
 Matches 13; Conservative 5; Mismatches 8; Indels 0; Gaps 0;
 QY 1341 YSEGVINKTVKMWIPALIQSNQSIL 1366
 DB 2907 HAFQVINNAYFLPALQOQSNNSNI 2932

RESULT 134
ID IMB_DROME STANDARD: PRI: 117 AA.
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE MITOCHONDRIAL IMPORT INNER MEMBRANE TRANSLOCASE SUBUNIT TIM9B.
GN IMB9 OR TIM9 OR C51757.
OS Irosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Eukaryota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Eukaryota; Insecta; Diptera; Eurytomidae; Irosophiidae; Irosophiinae;
OC NCBI_TaxID:7227;
RN [1]
RA SEQUENCE FROM N.A.
RA Burger M.F., Branner M., Hofmann S.,
RA "Cloning and mapping of the Tim9/Tim10 gene family encoding small zinc
RA finger proteins involved in mitochondrial carrier import";
RA Submitted (MAY 1999) to the EMBL/GenBank/Tran databases
RA [2]
RA SEQUENCE FROM N.A.
RA MELLINE BECKLEY;
RA MELLINE 2019006; PubMed 10731142;
RA Adams M.L., Scherzer S.F., Holt P.A., Evans C.A., Gecayene J.D.,
RA Aminiattides P.G., Scherzer S.F., Holt P.A., Evans C.A., Gecayene J.D.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Suttin G.G., Wortman J.K., Yandell M.D., Zhang G., Chen L.X.,
RA Branton R.G., Roberts Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Adhyan A., An H.-J., Andrews-Pfannkuch C., Baldwin P.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu I., Beasley E.M.,
RA Beeson K.Y., Bencs P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Bokorova E., Botchan M.F., Brack J., Brackstein P., Brothier P.,
RA Burtis K.C., Busan B.A., Butler H., Cadieu P., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Danke C., Davenport L.B., Davies P.,
RA DePallos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Deason K., Deep L.E., Downes M., Duman-Rocha S., Dunkov R.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferrara S., Ferrier A., Fleischmann W.,
RA Foster C., Gabriellian A.E., Gada N.S., Gelbart W.M., Glasser K.,
RA Glusok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.B., Harvey K., Helms T.S., Hernandez T.P., Hengst J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalili M., Kalish F., Karpen G.H., Ke Z., Kennison C.A., Ketchum K.A.,
RA Kimmel B.P., Kodira G., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Lin X., Maltsev B., McIntosh T.C., Meleod M.P., McPherson D.,
RA Metkalova D., Milshina N.V., Molnar C., Morris J., Moshireli A.,
RA Mount S.M., Moy M., Murphy R., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskorn D.P., Paclob J.M.,
RA Palazuelos M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Reinert K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Stiden-Kimble L., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A., Stappleton M., Strong P., Sun F.,
RA Sztraskas R., Teeter C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z., Wassenaar D.A., Weisslock G.M., Weisslock G.M.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zhou X.H., Zhou F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.G.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Irosophila melanogaster";
RA Science 287:2185-2195 (2000).
CC This SWISS PRO entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMRI outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.sib.ch/embassy>
CC or send an email to license@sib.ch).
EMBL: AF15104; AM43030.1;

DR EMBL: AE003512; AAF48985.1;
DR FlyBase: FBgn027358; Tim9.
KW Transport; Protein transport; Translocation; Mitochondrion;
KW Inner membrane.
SQ SEQUENCE 117 AA: 13520 MW: 3377A342075A57E5 CR664;
Query Match 0.5%; Score 58; DB 1; Length 117;
Best Local Similarity 20.0%; Pred. No. 26;
Matches 12; Conservative 18; Mismatches 30; Indels 0;
QY 1143 CAGIVSSVSGTGVNAEEVRIHFFHFAFPIVIVVVFQKKMGPKSGLDSVAFVWS 1202
DB 42 CVDRCVTKTAPFNQNMKVVDVWITINAKPMEEENAKKAEQGPQEKRLKAAAI 101
RESULT 134
ID VNSI_BRSVA STANDARD: PRI: 136 AA.
AC Q65694;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE NONSTRUCTURAL PROTEIN 1 (NONSTRUCTURAL PROTEIN 1C).
GN 1C OR NS1.
OS Bovine respiratory syncytial virus (strain A51908) (BRV).
OC Viruses; SSPNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus;
OC NCBI_TaxID=11247;
RN [1]
RA SEQUENCE FROM N.A.
RA MEDLINE-95146950; PubMed-7844532;
RA Pastey M.K., Samal S.K.;
RA "Nucleotide sequence analysis of the non-structural NS1 (1C) and NS2
RA (1B) proteins of bovine respiratory syncytial virus";
RA J. Gen. Virol. 76:193-197 (1995).
CC This SWISS-PRO entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMRI outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.sib.ch/embassy>
CC or send an email to license@sib.ch).
EMBL: U15947; AAA85671.1;
KW Nonstructural protein.
SQ SEQUENCE 136 AA: 15246 MW: AEB6A89D7BBD2D24 CR664;
Query Match 0.5%; Score 58; DB 1; Length 136;
Best Local Similarity 28.1%; Pred. No. 31;
Matches 16; Conservative 11; Mismatches 30; Indels 0;
QY 161 FVLLPVKQSCVPLAKETLTITHCYKGLCEMFPIQSLVTKSVKVFAPYDSSALPLVLL 217
DB 77 FVSMPLTQNSYVWELMELHCHPEQNTLHNNELIIFSSKPISSKFIKYSNGLSTLL 144
RESULT 135
ID IL2_MUOSSP STANDARD: PRI: 166 AA.
AC Q6R867; P70462; P70463;
DT 01-FEB-1995 (Rel. 31, Created)
DT 15-JUL-1995 (Rel. 38, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE INTERLEUKIN-2 PREPUSOR (IL-2) (T-CELL GROWTH FACTOR) (TCGF).
GN IL2.
OS Mus spretus (Western wild mouse).
OC Eukaryota; Metazoa; Chordata; Granata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OC NCBI_TaxID=10096;
RN [1]

01 15-JUL-1999 (rel. 60, last sequence update)
 02 15-JUL-1999 (rel. 60, last annotation update)
 03 FRAXALKINE PRECURSOR (NEUROKATIN) (CYTO MEMBRANE-ANCHORED)
 04 "HEM KINE" (SMALL INTRACELLULAR CYTOKINE)
 05 SVL1: P1EN: A011
 06 Ratius: 0000000000 (rel. 1)
 07 Fukuyama, M., et al. (1999) J. Neurochem. 69:1000-1008
 08 Mammalian Pathology: Robert J. Schumacher, Murinda; Ratius; Ratius;
 09 NBL: 0010101010
 10 SEQUENCE FROM N.A.
 11 TISSUE: Brain
 12 METLINE: 0000000000 (rel. 60, last sequence update)
 13 RA: 0000000000 (rel. 60, last sequence update)
 14 RA: 0000000000 (rel. 60, last sequence update)
 15 RA: 0000000000 (rel. 60, last sequence update)
 16 RA: 0000000000 (rel. 60, last sequence update)
 17 RA: 0000000000 (rel. 60, last sequence update)
 18 RA: 0000000000 (rel. 60, last sequence update)
 19 RA: 0000000000 (rel. 60, last sequence update)
 20 RA: 0000000000 (rel. 60, last sequence update)
 21 RA: 0000000000 (rel. 60, last sequence update)
 22 RA: 0000000000 (rel. 60, last sequence update)
 23 RA: 0000000000 (rel. 60, last sequence update)
 24 RA: 0000000000 (rel. 60, last sequence update)
 25 RA: 0000000000 (rel. 60, last sequence update)
 26 RA: 0000000000 (rel. 60, last sequence update)
 27 RA: 0000000000 (rel. 60, last sequence update)
 28 RA: 0000000000 (rel. 60, last sequence update)
 29 RA: 0000000000 (rel. 60, last sequence update)
 30 RA: 0000000000 (rel. 60, last sequence update)
 31 RA: 0000000000 (rel. 60, last sequence update)
 32 RA: 0000000000 (rel. 60, last sequence update)
 33 RA: 0000000000 (rel. 60, last sequence update)
 34 RA: 0000000000 (rel. 60, last sequence update)
 35 RA: 0000000000 (rel. 60, last sequence update)
 36 RA: 0000000000 (rel. 60, last sequence update)
 37 RA: 0000000000 (rel. 60, last sequence update)
 38 RA: 0000000000 (rel. 60, last sequence update)
 39 RA: 0000000000 (rel. 60, last sequence update)
 40 RA: 0000000000 (rel. 60, last sequence update)
 41 RA: 0000000000 (rel. 60, last sequence update)
 42 RA: 0000000000 (rel. 60, last sequence update)
 43 RA: 0000000000 (rel. 60, last sequence update)
 44 RA: 0000000000 (rel. 60, last sequence update)
 45 RA: 0000000000 (rel. 60, last sequence update)
 46 RA: 0000000000 (rel. 60, last sequence update)
 47 RA: 0000000000 (rel. 60, last sequence update)
 48 RA: 0000000000 (rel. 60, last sequence update)
 49 RA: 0000000000 (rel. 60, last sequence update)
 50 RA: 0000000000 (rel. 60, last sequence update)
 51 RA: 0000000000 (rel. 60, last sequence update)
 52 RA: 0000000000 (rel. 60, last sequence update)
 53 RA: 0000000000 (rel. 60, last sequence update)
 54 RA: 0000000000 (rel. 60, last sequence update)
 55 RA: 0000000000 (rel. 60, last sequence update)
 56 RA: 0000000000 (rel. 60, last sequence update)
 57 RA: 0000000000 (rel. 60, last sequence update)
 58 RA: 0000000000 (rel. 60, last sequence update)
 59 RA: 0000000000 (rel. 60, last sequence update)
 60 RA: 0000000000 (rel. 60, last sequence update)
 61 RA: 0000000000 (rel. 60, last sequence update)
 62 RA: 0000000000 (rel. 60, last sequence update)
 63 RA: 0000000000 (rel. 60, last sequence update)
 64 RA: 0000000000 (rel. 60, last sequence update)
 65 RA: 0000000000 (rel. 60, last sequence update)
 66 RA: 0000000000 (rel. 60, last sequence update)
 67 RA: 0000000000 (rel. 60, last sequence update)
 68 RA: 0000000000 (rel. 60, last sequence update)
 69 RA: 0000000000 (rel. 60, last sequence update)
 70 RA: 0000000000 (rel. 60, last sequence update)
 71 RA: 0000000000 (rel. 60, last sequence update)
 72 RA: 0000000000 (rel. 60, last sequence update)
 73 RA: 0000000000 (rel. 60, last sequence update)
 74 RA: 0000000000 (rel. 60, last sequence update)
 75 RA: 0000000000 (rel. 60, last sequence update)
 76 RA: 0000000000 (rel. 60, last sequence update)
 77 RA: 0000000000 (rel. 60, last sequence update)
 78 RA: 0000000000 (rel. 60, last sequence update)
 79 RA: 0000000000 (rel. 60, last sequence update)
 80 RA: 0000000000 (rel. 60, last sequence update)
 81 RA: 0000000000 (rel. 60, last sequence update)
 82 RA: 0000000000 (rel. 60, last sequence update)
 83 RA: 0000000000 (rel. 60, last sequence update)
 84 RA: 0000000000 (rel. 60, last sequence update)
 85 RA: 0000000000 (rel. 60, last sequence update)
 86 RA: 0000000000 (rel. 60, last sequence update)
 87 RA: 0000000000 (rel. 60, last sequence update)
 88 RA: 0000000000 (rel. 60, last sequence update)
 89 RA: 0000000000 (rel. 60, last sequence update)
 90 RA: 0000000000 (rel. 60, last sequence update)
 91 RA: 0000000000 (rel. 60, last sequence update)
 92 RA: 0000000000 (rel. 60, last sequence update)
 93 RA: 0000000000 (rel. 60, last sequence update)
 94 RA: 0000000000 (rel. 60, last sequence update)
 95 RA: 0000000000 (rel. 60, last sequence update)
 96 RA: 0000000000 (rel. 60, last sequence update)
 97 RA: 0000000000 (rel. 60, last sequence update)
 98 RA: 0000000000 (rel. 60, last sequence update)
 99 RA: 0000000000 (rel. 60, last sequence update)
 100 RA: 0000000000 (rel. 60, last sequence update)

RESULT 146
 QX1: SWIN
 ID: QX1: SWIN
 AC: Q02226
 DT: 01-07-1999 (rel. 27, last sequence update)
 DT: 01-07-1999 (rel. 27, last sequence update)
 DT: 30-MAY-2000 (rel. 39, last annotation update)
 DE: CYTOCHROME C OXIDASE POLYPEPTIDE 11 PRECURSOR (EC 1.9.3.1) (FRAGMENT)
 GN: COX2
 QS: glycolytic max (Soybean)
 OC: Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta
 OC: Monilophyta; Eudicotyledons; Core eudicot; Rosidae; Euphorbiales
 OC: Euphorbiales; Fabaceae; Papilionaceae; Glycyne
 OC: NCBI_TaxID 3847
 RN: [1]
 RP: SEQUENCE FROM N.A.
 RC: STRAIN: CV, WILLIAMS; TISSUE: S root, and Leaf;
 RX: MEDLINE: 94010475; PubMed: 1362479;
 RA: Corallo P.S., Gray M.W.
 RT: "Silent mitochondrial and active nuclear genes for subunit 2 of
 cytochrome c oxidase (cox2) in soybean: evidence for RNA-mediated
 gene transfer."
 RL: EMBO J. 13:3815-3820(1992).
 CC: 1- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY
 CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNIT 2
 3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. SUBUNIT 2
 TRANSFERS THE ELECTRONS FROM CYTOCHROME C VIA ITS HEMICENTRAL COPPER
 A CENTER TO THE HEMICENTRAL COPPER A CENTER OF THE CATALYTIC SUBUNIT 1.
 CC: 1 CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C OXIDASE [EC 1.9.3.1]
 CC: 4 FERROCYTOCHROME C
 CC: 1 SUBCELLULAR LOCATION: INTERNAL MEMBRANE PROTEIN. MITOCHONDRIAL
 INNER MEMBRANE.
 CC: 1 SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 2 FAMILY.
 CC: This SWISS-Prot entry is copy right. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non profit institutions as long as its content is in no way
 modified and this statement is not removed, misuse by and for commercial
 entities requires a license agreement (see http://www.ebi.ac.uk/seqdb/doc/seqdb.html
 or send an email to license@ebi.ac.uk).
 DR: EMBL: Z11980; CAA78032.1;
 DR: F1R: S28027; S28027.
 DR: HSP: P00404; I000.
 DR: M0004; 13688; GYMMA_C0A222.
 DR: InterPro: IPR001505;
 DR: InterPro: IPR002429;
 DR: Pfam: PF00114; COX2; 1;
 DR: ProSITE: PS00078; COX2; 1;
 KW: oxidoreductase; Copper; Mitochondrion; Transmembrane; Inner membrane;
 Electron transport; Respiratory chain; Transit peptide.
 FT: NON_TER 1
 FT: TRANSIT 1 119
 FT: CHAIN 2120 394
 FT: DOMAIN 150 181
 FT: TRANSMEM 182 198
 FT: DOMAIN 199 226
 FT: DOMAIN 234 244
 FT: DOMAIN 247 254
 FT: METAL 462 462
 FT: METAL 466 466
 FT: METAL 470 470
 SQ: SEQUENCE 394 AA; 44334 MW; CDE7AE60DAE530B C0664;
 Query Match: 0.58; Score 58; DB 1; Length 394;
 Best Local Similarity 40.58; Pred. No. 1; Length 394;
 Matches 15; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

1000

[2] SEQUENCE FROM N.A.
 RP STRAIN SEROVAR 4
 KC MEDLINE 29500219; PubMed 11048724;
 RA Glass J.E., Leikowitz E.J., Glass J.E., Heiner C.P., Chen E.Y.,
 RA Cassell G.H.;
 RT "The complete sequence of the mucosal pathogen *Ureaplasma*
 RT *urealyticum*.";
 RL Nature 407:757-762(2000).
 RN [3]
 CC SEQUENCE OF 1-135 FROM N.A.
 CC STRAIN VARIANTS SEROVARS
 CC MEDLINE 29023972; PubMed 10555722;
 RA Kong F., James G., Ma H., Gordon S., Kwak B., Gilbert G.L.;
 RT "Phylogenetic analysis of *Ureaplasma urealyticum* -- support for the
 RT establishment of a new species, *Ureaplasma parvum*.";
 RL Int. J. Syst. Bacteriol. 49:1879-1889(1999).
 CC -1- CATALYTIC ACTIVITY: GREA + H(2O) -> CO(2) + 2 NH(3).
 CC -1- COFACTOR: EACH CHAIN BINDS TWO NICKEL IONS.
 CC -1- SUBUNIT: ALPHA, BETA, GAMMA(3) (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE UREASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: L40489; AAC89189.2; -;
 CC EMBL: AF002140; AAF36944.1; -;
 CC EMBL: AF085730; AAD28136.1; -;
 CC EMBL: AF085731; AAD28139.1; -;
 CC EMBL: AF085732; AAD28142.1; -;
 CC EMBL: AF085733; AAD28145.1; -;
 CC HSSP: P18314; 1PWE.
 CC InterPro: IPR001924.
 CC Pfam: PF00449; urease; 1.
 CC PROSITE: PS01120; UREASE_1; 1.
 CC PROSITE: PS00145; UREASE_2; 1.
 CC Hydrolase; Metal-binding; Nickel.
 CC METAL 141 141 NICKEL 2 (BY SIMILARITY).
 CC METAL 143 143 NICKEL 2 (BY SIMILARITY).
 CC METAL 223 223 NICKEL 1 AND 2 (BY SIMILARITY).
 CC METAL 252 252 NICKEL 1 (BY SIMILARITY).
 CC METAL 278 278 NICKEL 1 (BY SIMILARITY).
 CC METAL 366 366 NICKEL 2 (BY SIMILARITY).
 CC ACT_SITE 326 326 BY SIMILARITY.
 CC CONFLICT 100 100 I -> L (IN REF. 1).
 CC CONFLICT 355 355 H -> A (IN REF. 1).
 CC SEQUENCE 598 AA: 64544 MW: 57601366021424 CRC64;
 Query Match 0.5%; Score 58; DB 1; Length 598;
 Best Local Similarity 43.7%; Pred. No. 1,9c-02;
 Matches 11; Conservative 7; Mismatches 9; Indels 0; Gaps 0;
 QY 670 DNINLGHSSMKWVDFISVGDESP 696
 LB 19 DSVRLGDTNWKVKVKKLLVNGLSVF 45
 RESULT 160
 VP4_ROT15 STANDARD; PRT; 775 AA.
 AC P40214;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DE OUTER CAPSID PROTEIN VP4 (HEMAGGLUTININ) (OUTER LAYER PROTEIN VP4)
 DE [CONTAINS: OUTER CAPSID PROTEINS VP5 AND VP8].
 GN S4.

OS Human rotavirus (serotype 2 / strain RV-5).
 OC Viruses; dsRNA viruses; reoviridae; Rotavirus.
 OX NCBI_TaxID=31569;
 RN [1]
 RP SEQUENCE FROM N.A.
 KC MEDLINE=87127560; PubMed=3028337;
 RA Kantharidis P., Drall-Smith M.L., Holmes I.H.;
 RT "Marked sequence variation between segment 4 genes of human RV-5 and
 RT simian SA 11 rotaviruses.";
 RL Arch. Virol. 93:111-121(1987).
 CC -1- SUBCELLULAR LOCATION: OUTER CAPSID.
 CC -1- PTM: VP8 IS ONE OF TWO TRYPSIN CLEAVAGE PRODUCTS OF VP4; THE
 CC OTHER PRODUCT IS VP5.
 CC -1- SIMILARITY: BELONGS TO THE ROTAVIRUS VP4 PROTEIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M32559; AAA47333.1; -;
 CC PIP: A43760; VPXPMF
 CC InterPro: IPR000416; -;
 CC Pfam: PF00426; VP4; 1
 CC Coat protein; glycoprotein.
 CC CHAIN 1 775 OUTER CAPSID PROTEIN VP4.
 CC CHAIN 1 240 OUTER CAPSID PROTEIN VP8.
 CC CAPROHYD 56 56 N-LINKED (GLCNAC...) (POTENTIAL).
 CC CARBOHYD 132 132 N-LINKED (GLCNAC...) (POTENTIAL).
 CC CARBOHYD 195 195 N-LINKED (GLCNAC...) (POTENTIAL).
 CC CARBOHYD 324 324 N-LINKED (GLCNAC...) (POTENTIAL).
 CC CARBOHYD 583 583 N-LINKED (GLCNAC...) (POTENTIAL).
 CC CARBOHYD 589 589 N-LINKED (GLCNAC...) (POTENTIAL).
 CC CAPROHYD 592 592 N-LINKED (GLCNAC...) (POTENTIAL).
 CC CARBOHYD 599 599 N-LINKED (GLCNAC...) (POTENTIAL).
 CC SEQUENCE 775 AA: 87593 MW: 86760868267186 CRC64;
 Query Match 0.5%; Score 58; DB 1; Length 775;
 Best Local Similarity 45.6%; Pred. No. 2,6a-02;
 Matches 16; Conservative 7; Mismatches 22; Indels 0; Gaps 0;
 QY 1045 IQKPTAVIKKEMAVIHITQYNEFSVSLINEDPKSPIDPIKAV 1089
 LB 667 IPKSYVVLKINFMVEMINTEKPFAYKVDLINEIPFDINKFAELV 711
 RESULT 161
 VP4_ROT1L STANDARD; PRT; 775 AA.
 AC P21284;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)
 DE OUTER CAPSID PROTEIN VP4 (HEMAGGLUTININ) (OUTER LAYER PROTEIN VP4)
 DE [CONTAINS: OUTER CAPSID PROTEINS VP5 AND VP8].
 GN S4.
 OS Human rotavirus (strain 126).
 OC Viruses; dsRNA viruses; reoviridae; Rotavirus.
 OX NCBI_TaxID=10953;
 RN [1]
 RP SEQUENCE FROM N.A.
 KC MEDLINE=91012813; PubMed=2170690;
 RA Taniguchi K., Urasawa T., Kobayashi N., Gorziglia M., Urasawa S.;
 RT "Nucleotide sequence of VP4 and VP7 genes of human rotaviruses with
 RT subgroup 1 specificity and long RNA pattern: implication for new G
 RT serotype specificity.";
 RL Virol. 64:5640-5644(1990).
 CC -1- SUBCELLULAR LOCATION: OUTER CAPSID.

11 1 CMC VIB IS ONE OF THE MYOSIN LEAVAGE PRODUCTS OF VI4; THE
 12 THE P-LEAVAGE IS VIB.
 13 1 SIMILARITY: HIGH, TO MYOSIN HEAVY CHAINS.
 14
 15 THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 16 between the Swiss Institute of Bioinformatics and the EMBL outstation
 17 at the European Bioinformatics Institute. There are no restrictions on its
 18 use by non-profit institutions as long as its content is in no way
 19 modified and this statement is not removed. Usage by and for commercial
 20 entities requires a license agreement (see <http://www.ebi.ac.uk/edm/>)
 21 or send an email to license@ebi.ac.uk.
 22
 23 EMBL: M62922; AAA17365.1
 24 PIR: A04100; VPXHWI
 25 PIR: A04100; VPXHWI
 26 UniProt: P00041; VPXHWI
 27 UniProt: P00041; VPXHWI
 28 UniProt: P00041; VPXHWI
 29 UniProt: P00041; VPXHWI
 30 UniProt: P00041; VPXHWI
 31 UniProt: P00041; VPXHWI
 32 UniProt: P00041; VPXHWI
 33 UniProt: P00041; VPXHWI
 34 UniProt: P00041; VPXHWI
 35 UniProt: P00041; VPXHWI
 36 UniProt: P00041; VPXHWI
 37 UniProt: P00041; VPXHWI
 38 UniProt: P00041; VPXHWI
 39 UniProt: P00041; VPXHWI
 40 UniProt: P00041; VPXHWI
 41 UniProt: P00041; VPXHWI
 42 UniProt: P00041; VPXHWI
 43 UniProt: P00041; VPXHWI
 44 UniProt: P00041; VPXHWI
 45 UniProt: P00041; VPXHWI
 46 UniProt: P00041; VPXHWI
 47 UniProt: P00041; VPXHWI
 48 UniProt: P00041; VPXHWI
 49 UniProt: P00041; VPXHWI
 50 UniProt: P00041; VPXHWI
 51 UniProt: P00041; VPXHWI
 52 UniProt: P00041; VPXHWI
 53 UniProt: P00041; VPXHWI
 54 UniProt: P00041; VPXHWI
 55 UniProt: P00041; VPXHWI
 56 UniProt: P00041; VPXHWI
 57 UniProt: P00041; VPXHWI
 58 UniProt: P00041; VPXHWI
 59 UniProt: P00041; VPXHWI
 60 UniProt: P00041; VPXHWI
 61 UniProt: P00041; VPXHWI
 62 UniProt: P00041; VPXHWI
 63 UniProt: P00041; VPXHWI
 64 UniProt: P00041; VPXHWI
 65 UniProt: P00041; VPXHWI
 66 UniProt: P00041; VPXHWI
 67 UniProt: P00041; VPXHWI
 68 UniProt: P00041; VPXHWI
 69 UniProt: P00041; VPXHWI
 70 UniProt: P00041; VPXHWI
 71 UniProt: P00041; VPXHWI
 72 UniProt: P00041; VPXHWI
 73 UniProt: P00041; VPXHWI
 74 UniProt: P00041; VPXHWI
 75 UniProt: P00041; VPXHWI
 76 UniProt: P00041; VPXHWI
 77 UniProt: P00041; VPXHWI
 78 UniProt: P00041; VPXHWI
 79 UniProt: P00041; VPXHWI
 80 UniProt: P00041; VPXHWI
 81 UniProt: P00041; VPXHWI
 82 UniProt: P00041; VPXHWI
 83 UniProt: P00041; VPXHWI
 84 UniProt: P00041; VPXHWI
 85 UniProt: P00041; VPXHWI
 86 UniProt: P00041; VPXHWI
 87 UniProt: P00041; VPXHWI
 88 UniProt: P00041; VPXHWI
 89 UniProt: P00041; VPXHWI
 90 UniProt: P00041; VPXHWI
 91 UniProt: P00041; VPXHWI
 92 UniProt: P00041; VPXHWI
 93 UniProt: P00041; VPXHWI
 94 UniProt: P00041; VPXHWI
 95 UniProt: P00041; VPXHWI
 96 UniProt: P00041; VPXHWI
 97 UniProt: P00041; VPXHWI
 98 UniProt: P00041; VPXHWI
 99 UniProt: P00041; VPXHWI
 100 UniProt: P00041; VPXHWI

RP PHOSPHORYLATION.
 RX MEDLINE:8942906; PubMed:2754744;
 RA Schriber L.A.; Waterson R.H.;
 RT "Phosphorylation of the N-terminal region of Cerezo-habitat is essential
 RL J. Mol. Biol. 207:451-454 (1989).
 CC 1 FUNCTION: PARAMYOSIN IS A MAJOR STRUCTURAL COMPONENT OF
 CC MANY THICK FILAMENTS ISOLATED FROM INVERTEBRATE MUSCLES.
 CC 1 SURNAME: HOMODIMER (HY SIMILARITY).
 CC 1 DOMAIN: FOR MOST OF ITS LENGTH, PARAMYOSIN APPEARS TO FORM AN
 CC ALPHA-HELICAL COILED COIL AND SHOWS THE HEPTAD REPEAT OF
 CC HYDROPHOBIC AMINO ACID RESIDUES AND THE 28-RESIDUE REPEAT OF
 CC CHARGED AMINO ACIDS CHARACTERISTIC OF MYOSIN HEAVY CHAINS.
 CC HOWEVER, PARAMYOSIN DIFFERS FROM MYOSIN IN HAVING NON-HELICAL
 CC EXTENSIONS AT BOTH TERMINI AND AN ADDITIONAL "SKIP" RESIDUE THAT
 CC INTERRUPTS THE 28-RESIDUE REPEAT. THE DISTRIBUTION OF CHARGED
 CC RESIDUES IS ALSO DIFFERENT FROM MYOSIN HEAVY CHAINS.
 CC 1 PTM: PHOSPHORYLATED ON SERINE RESIDUES IN THE N-ALPHA-HELICAL
 CC N-TERMINAL REGION.
 CC 1 SIMILARITY: HIGH, TO MYOSIN HEAVY CHAINS.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/edm/>)
 CC or send an email to license@ebi.ac.uk.
 CC
 CC EMBL: X08098; CAA10857.1;
 CC PIR: 504027; S04027;
 CC RCSB: 180220; 10318;
 CC InterPro: IPR002928;
 CC Pfam: PF01576; Myosin_tail; 1;
 CC Coiled coil; Heptad repeat pattern; Muscle protein; Thick filament;
 KW Myosin; Phosphorylation.
 FT DOMAIN 1 47 NON-HELICAL REGION (POTENTIAL).
 FT DOMAIN 48 871 COILED COIL (POTENTIAL).
 FT DOMAIN 872 882 NON-HELICAL REGION (POTENTIAL).
 FT DISULFID 143 143 INTERCHAIN (POTENTIAL).
 FT CLEAVAGE 532 532 SILENT (POTENTIAL).
 SQ SEQUENCE 882 AA; 101949 MW; 89404846705F CRO64;
 Query Match 0.58; Score 58; DB 1; Length 882;
 Best Local Similarity 23.88; Pred. NG: 30-02;
 Matches 20; Conservative 13; Mismatches 54; Indels 6; Gaps 0;
 QY 1732 QLPSTLMSESLIMKNSISVSSVLSALALQKVVLLHPIISYVGLISQVILK 1741
 DD 593 QUKLQASLHDTQKQVQVLDVYALALAVAAI SALLERKALALAKAVAKKNEVLE 652
 QY 1792 ITSEMGASQANPLTSIKKTLATILA 1818
 DB 653 ANGRISDLISINNLTSLKLETELS 679
 RESULT 163
 ID UP0_HUMAN STANDARD; PRT; 687 AA.
 AC P40530;
 DT 01-APR-1993 (rel. 25, Created)
 DT 01-APR-1993 (rel. 25, Last sequence update)
 DT 01-OCT-2000 (rel. 40, Last annotation update)
 DE TYPE-SINE-PROTEIN KINASE RECEPTOR (EC 2.7.1.112) (AXL
 DE ONCOGENE).
 GN AXL OR UP0.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID:9606;
 RN [1]

SEQUENCE FROM N.A.
MEDLINE: 92050809; PubMed-1834974;
Jansson J.W.G., Schulz A.S., Steenvoorden A.C.M., Schmidberger M.,
Strobel S., Ambros P., Bartram C.R.;
"A novel putative tyrosine kinase receptor with oncogenic potential.";
Oncogene 6:2113-2120(1991).
[2]
SEQUENCE FROM N.A.
MEDLINE: 92017777; PubMed-1656226;
O'Bryan J.P., Frye R.A., Codswell P.C., Neubauer A., Kitch B.,
Prokop C., Espinosa R., le Beau M.M., Earp H., Liu E.T.;
"AXL, a transforming gene isolated from primary human myeloid
leukemia cells, encodes a novel receptor tyrosine kinase.";
Mol. Cell. Biol. 11:5016-5031(1991).
[4]
SEQUENCE OF 667-723 FROM N.A.
MEDLINE: 94067791; PubMed-9247543;
Lee S.-T., Strunk K.M., Spritz K.A.;
"A survey of protein tyrosine kinase mRNAs expressed in normal human
melanocytes.";
Oncogene 8:3403-3410(1993).
-1- FUNCTION: MAY FUNCTION AS A SIGNAL TRANSDUCER BETWEEN SPECIFIC
CELL TYPES OF MESODERMAL ORIGIN.
-2- CATALYTIC ACTIVITY: ATP, A PROTEIN TYROSINE - ADP, PROTEIN
TYROSINE PHOSPHATE.
-3- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
-4- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A
SHORT FORM ARE PRODUCED BY ALTERNATIVE SPLICING.
-5- DISEASE: HAS TRANSFORMING POTENTIAL IN PATIENTS WITH CHRONIC
MYELOPROLIFERATIVE DISORDER OR CHRONIC MYELOIDIC LEUKEMIA.
-6- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.
-7- SIMILARITY: CONTAINS 2 IMMUGLOBULIN LIKE C2 TYPE DOMAINS.
-8- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III LIKE DOMAINS.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/annou>
or send an email to license@isb-sib.ch).

EMBL: X57019; CAA0338.1; ALT_INIT.
EMBL: M76125; AAA61243.1; ALT_INIT.
EMBL: S65125; AAB20305.1; ALT_INIT.
HSSP: P11362; IFGI.
MIM: 109135;
InterPro: IPR000719;
InterPro: IPR001245;
InterPro: IPR001777;
InterPro: IPR003006;
Pfam: PF00041; I03; 2;
Pfam: PF00047; I03; 2;
Pfam: PF00069; pkinase; 1;
PROSITE: PS00107; PROTEIN_KINASE_ATP; 1;
PROSITE: PS00109; PROTEIN_KINASE_TYR; 1;
PROSITE: PS00011; PROTEIN_KINASE_DOM; 1;
Receptor; Glycoprotein; Tyrosine kinase, ATP binding;
transferase; phosphorylation, Transmembrane, Signal; Repeat;
Immunoglobulin domain; Proto-oncogene; Alternative splicing.
POTENTIAL.
FT CHAIN 1 18
FT SIGNAL 1 18
FT CHAIN 19 887
FT DDOMAIN 19 444
FT TRANSMEM 445 465
FT CYTOPLASMIC 466 887
FT DDOMAIN 466 887
FT DDOMAIN 42 117
FT DDOMAIN 146 205
FT DDOMAIN 217 315
FT DDOMAIN 319 416
FT DDOMAIN 529 600
FT DDOMAIN 535 543
FT NP_BIND 543 543
FT BINDING 560 560
ATP (BY SIMILARITY).
ATP (BY SIMILARITY).

FT ACT_SITE 665 665 BY SIMILARITY.
FT DISULFID 49 110 BY SIMILARITY.
FT DISULFID 153 198 BY SIMILARITY.
FT CARBOHYD 36 36 N-LINKED (GLNAC. . .) (POTENTIAL).
FT CARBOHYD 150 150 N-LINKED (GLNAC. . .) (POTENTIAL).
FT CARBOHYD 191 191 N-LINKED (GLNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLNAC. . .) (POTENTIAL).
FT CARBOHYD 338 338 N-LINKED (GLNAC. . .) (POTENTIAL).
FT CARBOHYD 394 394 N-LINKED (GLNAC. . .) (POTENTIAL).
FT MOD_RES 496 496 PHOSPHORYLATION (AUTO.) (BY SIMILARITY).
FT VARSPIC 422 430 MISSING (IN SHORT ISOFORM).
FT CONFLICT 235 236 P > T (IN REF. 2).
FT CONFLICT 331 331 E > K (IN REF. 2).
FT CONFLICT 632 632 G > D (IN REF. 2).
SQ SEQUENCE 887 AA; 97374 MW; 24/RECI646298EDA CKB04;
Query Match 0.5%; Score 58; DR 1; Length 887;
Best Local Similarity 25.0%; From RV 3502;
Matches 15; Conservative 15; Mismatches 30; Indels 0; Gaps 0;
CY 1216 LLEGLQHKKELESFHLVPT;EHLSEF;FHLSEFQFQFHEHYTVAIIISGINTQV;SPS 1269
DB 751 IYNYLPQCNELKQAPDLGLIYALMSRWELNPDQRFSTFELREDLENTLKALPPAQEPD 810
RESULT 164
IF2_CHAIN STANVAPD PPT: 892 AA.
AC 084098:
DT 30-MAY-2000 (Ref. 39, Created)
DT 30-MAY-2000 (Ref. 39, Last sequence update)
DT 30-MAY-2000 (Ref. 39, Last annotation update)
DE TRANSLATION INITIATION FACTOR IF-2.
GN INEB OR CT096.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
PP SEQUENCE FROM N.A
FC STEFAN T. VAV 3/98;
EX MPMI:PMF-94000809; PubMed=9784136;
PA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe P., Aravind L.,
PA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
PA Davis R.W.;
RA "Genome sequence of an obligate intracellular pathogen of humans:
Chlamydia trachomatis.";
RT Chlamydia trachomatis.";
RL Science 282:754-759(1998).
CC -1- FUNCTION: IF-2, ONE OF THE ESSENTIAL COMPONENTS FOR THE INITIATION
OF PROTEIN SYNTHESIS IN VITRO, PROTECTS FORMYETHIONYL-TRNA FROM
SPONTANEOUS HYDROLYSIS AND PROMOTES ITS BINDING TO THE 30S
RIBOSOMAL SUBUNIT. IT IS ALSO INVOLVED IN THE HYDROLYSIS OF GTP
DURING THE FORMATION OF THE 70S RIBOSOMAL COMPLEX (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE IF-2 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/annou>
or send an email to license@isb-sib.ch).

EMBL: AE001283; AAC67687.1;
InterPro: IPR000178;
InterPro: IPR000795;
Pfam: PF00009; GTP_EFTD; 1;
Pfam: PF02131; IF2; 1;
PROSITE: PS01176; IF2; 1;
Initiation factor; Protein biosynthesis; GTP-binding.
FT DDOMAIN 400 548
FT NP_BIND 406 413
GTP (BY SIMILARITY).

FT	I=MAIN	655	960	CELLULOSE BINDING (BY SIMILARITY).
FT	AUT SITE	450	560	PROTEIN BINDER (BY SIMILARITY).
FT	AUT SITE	565	565	RNASE-HLIE (BY SIMILARITY).
FT	I=MAIN	833	895	2 X 24 AA APPROXIMATE REPEATS.
FT	REPEAT	833	856	
FT	REPEAT	872	895	
FT	SERULINIC	900 AA:	102415 MW	Z: 97.9AHL1954F024-03-SP64:

Query Match	0.5%	Score 5%	Id 1%	Length 50%
Best Local Similarity	22.8%	Prod. No. 4,110-02		
Matches	14	Mismatches	49	Index
	Conservative			50%

[illegible]

HISTONE TIA ON SERINE AND THREONINE RESIDUES (BY SOLUBILITY).
 -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN LYMPHOID ORGANS.
 -1- PTM: AUTOPHOSPHORYLATED (BY SIMILARITY).
 -1- SIMILARITY: BELONGS TO THE SEVENTH FAMILY OF PROTEIN KINASES.
 SIZE20 SUBFAMILY.
 THIS WORK WAS CARRIED OUT IN CO-OPERATION WITH THE EMBO OUTSTATION
 between the Swiss Institute of Bioinformatics and the EMBO outstation

use by non-profit institutions as long as its content is in no way
the European Bioinformatics Institute, there are no restrictions on its
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb.scrib.ac.uk/embnet/>
or send an email to license@isb-scrib.ch).

EMBL; AB015718; BAA35073; 1; -
EMBL; AL143081; CAB61400.1; -
HSSP; P24941; 1HCL.
MIM; 603919; -;
InterPro; IPR000719; -;
InterPro; IPR001245; -;
InterPro; IPR002290; -;
Pfam; PF00069; pkinase; 1;
PRINTS; PR00169; TYRKINASE;
PROSITE; PS00107; PROTEIN_KINASE_A1P; 1;
PROSITE; PS00109; PROTEIN_KINASE_SH; 1;

RC TISSUE: Brain;
RA Lopez-Rodriguez C., Avasthiani F., Pakeman A.S., Copeland N.G.,
RT Gilbert D.J., Thomas S., Distelhoe C., Jenkins N.A., Rao A.;
RT "RATS: The Rf-AT family of transduction factors expands in a new
RT direction.";
RT Cold Spring Harb. Symp. Quant. Biol. 64:517-526(1999).
RL [5].
RN SEQUENCE OF 675 1531 FROM N.A.
RP TISSUE: Brain;
RX MEDLINE 2002/2/28; PubMed 10765528;
KA Eschke C., Kell E., Jahnsmeyer A., Grosschik E.H., Schwinger E.;
RT "Isolation and characterization of novel CAG repeat containing genes
RT expressed in human brain.";
RL DNA Seq. 10:1-6(1999).
CC -1- FUNCTION: PLAYS A ROLE IN THE REGULATIVE EXPRESSION OF GENES.
CC REGULATES HYPERIONICALLY INDUCED CELLULAR ACCUMULATION OF
CC OSMOLYTES.
CC -1- SUBUNIT: DOES NOT BIND WITH PDS AND JUN TRANSCRIPTION FACTORS. BUT
CC MIGHT BE CAPABLE OF FORMING STABLE DIMERS WITH DNA ELEMENTS.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- ALTERNATIVE SPLICING: 3 ISOFORMS: A, B AND C (SHOWN HERE); MAY BE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: HIGHEST LEVELS IN SKELETAL MUSCLE, BRAIN,
CC HEART AND PERIPHERAL BLOOD LEUKOCYTES. ALSO EXPRESSED IN PLACENTA,
CC LUNG, LIVER, KIDNEY, PANCREAS, SPLEEN, THYMUS, PROSTATE, TESTIS,
CC OVARY, SMALL INTESTINE AND COLON.
CC -1- SIMILARITY: BELONGS TO THE REL/DORSAL FAMILY.
CC -1- CAUTION: REF.5 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 1165
CC ONWARD DUE TO A FRAMESHIFT.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: AB020634; BAA74850.1; -;
CC EMBL: AF089824; AAD18136.1; -;
CC EMBL: AF134870; AAD38360.1; -;
CC EMBL: Z97016; CA009693.1; ALT_FRAME.
CC EMBL: AF163836; AAD48441.1; -;
CC MIM: 604708; -;
CC InterPro: IPR000451; -;
CC PROSITE: PS01264; REL.1; PALM_NRG.
CC PROSITE: PS0254; REL.2; 1.
CC Transcription regulation; Activator; Nuclear protein; DNA-binding;
KW Alternative splicing.
FT DOMAIN 69 100 SER-RICH.
FT DOMAIN 293 350 GRA BINDING.
FT DOMAIN 739 743 POLY GLR.
FT DOMAIN 879 888 POLY-GLN.
FT DOMAIN 966 971 POLY-THR.
FT DOMAIN 1248 1266 POLY GLR.
FT VARSPIC 1 76 MISSING (IN ISOFORM A).
FT VARSPIC 1 47 MISSING (IN ISOFORM B).
FT C-REFLECT 1369 1369 E > D (IN REF. 5).
SQ SEQUENCE 1531 AA; 165764 MW; A686560880A8F69E CR664;

Query Match 0.5%; Score 58; DB 1; Length 1531;
Best Local Similarity 26.4%; Pred. No. 60+03;
Matches 19; Conservative 14; Mismatches 36; Indels 6; Gaps 0;
QY 123 CHLDGSGSLATVITHTETLIVPVVQLKFKRKEHWEHILFVPSQVPLAKPTLITTC 192
DB 496 MELPHQHLIVKVPYHQHITLIPVSVGLIVVVTNAGSHDVGPFYTYTIDPAAAGALNVN 555
QY 184 YKGLGFMDFICS 194
DB 556 KKEISSPAKICS 567

RESULT 175
POLG_POL2W STANDARD PPT: 2205 AA.
ID POLG_POL2W
AC P23069;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Pol 20, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: COAT PROTEIN VP1 TO VP4; CORE PROTEINS
DE P2A TO P2C, P3A; GENOME-LINKED PROTEIN VP3; PICORNAIN 3C
DE (PC 3.4.22.28) (PROTEASE 3C) (P3C); RNA-DIRECTED RNA POLYMERASE P40
DE (EC 2.7.7.48)].
OS Poliovirus type 2 (strain W-2).
OC Viruses; ssRNA positive strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus
OX NCBI_TaxID:12085;
RN [1]
FP SEQUENCE FROM N A
FX MEDLINE 90155230; PubMed 2154539;
FA Peavear D.C., Oh C.K., Cunningham L.L., Calenoff M., Jubelt R.;
RT "Localization of genomic regions specific for the attenuated, mouse-
RT adapted poliovirus type 2 strain W-2.";
PL J. Gen. Virol. 71:43-52(1990).
CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC Q/C SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS.
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: D00625; XA000516.1; ALT_SEQ.
CC FTR: A34832; GNRX2W.
CC HSP: P04299; IPOV.
CC MSOPS: C03.001; -;
CC MSOPS: C03.020; -;
CC InterPro: IPR000081; -;
CC InterPro: IPR000149; -;
CC InterPro: IPR000605; -;
CC InterPro: IPR001205; -;
CC InterPro: IPR001676; -;
CC InterPro: IPR002527; -;
CC Pfam: PF00548; cys_protease-3C; 1.
CC Pfam: PF00947; plico_P2A; 1.
CC Pfam: PF01552; plico_P2B; 1.
CC Pfam: PF00680; RNA_dep_RNA_pol; 1.
CC Pfam: PF00910; RNA_helicase; 1.
CC Pfam: PF00073; rhv; 3.
CC Polyprotein; Core protein; Transferase;
KW RNA directed RNA polymerase, Polyprotein; Myristate.
FT CHAIN 2 69
FT CHAIN 70 340
FT CHAIN 341 578
FT CHAIN 579 879
FT CHAIN 880 1028
FT CHAIN 1029 1125
FT CHAIN 1126 1454
FT CHAIN 1455 1541
FT CHAIN 1542 1563
FT CHAIN 1564 1746
FT CHAIN 1747 2205
FT LIPID 2 2
FT ACT_SITE 1710 1710
FT ACT_SITE 1724 1724
FT ACT_SITE 2205 AA; 245703 MW; 2A42AB039F0254AD CR664;
SQ SEQUENCE

OX NCBI_TaxID=2190;
 [1]
 RP SEQUENCE FROM N.A.
 KC STRAIN-1AL_1 / GSM 2661 / ATCC 43067,
 RX MEDLINE 96337904; PubMed 8624027;
 RA Buit C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Saiton G.G., Blake T.A., Fitzgerald T.M., Clayton P.A., Gocayne J.D.,
 RA Kariya A.P., Dougherty R.A., Tomb J.F., Adams M.F., Belch C.L.,
 RA Overbeck R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glisak A.,
 RA Scott J.L., Georger N.S.M., Weidman J.P., Fuhmann J.L., Nguyen D.,
 RA Ulterback T.P., Kelley T.M., Peterson S.N., Rhee P.W., Hwang M.C.,
 RA Clifton M.D., Roberts K.M., Hurst M.A., Kaine R.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.G., Weiss C.R., Venter J.C.;
 RI "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RI jannaschii.";
 RL Science 273:1058-1073(1996).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.1st-sib.ch/announcements/>
 CC or send an email to 1st-sib.ch@sib.ch).
 CC EMBL: 177118; AAC37103.1; -
 DR TIGR: M10734; -
 KW Hypothetical protein.
 SQ SEQUENCE 295 AA; 33623 MW; 191PTA001002151 CRF64;

Query Match 0.54; Score 57; DB 1; Length 295;
 Best Local Similarity 32.39; Pred. No. 1,1e+02;
 Matches 12; Conservative 8; Mismatches 16; Indels 0; Gaps 0.
 QY 1625 RTGCGHSNPKTAVFELKLVLAALVQREKKE 1660
 Db 260 RTGSPGTFETINFLKPLKLVLAALVQREKRE 295

RESULT 189
 GCP_MYCGE
 ID GCP_MYCGE STANDARD; PRT; 315 AA.
 AC P47292;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PROBABLE G-STALOGLYCOPROTEIN ENDOPEPTIDASE (EC 3.4.24.57)
 DE (GLYCOPROTEASE).
 GN GCP OR M0046.
 OS Mycoplasma genitalium
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 OC Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2097;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC SEPAIN AF033530 / G 27;
 RX MEDLINE 96326446; PubMed 7569993;
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton P.A.,
 RA Fleischmann R.D., Holt C.E., Kariya A.P., Saiton G., Kelley T.M.,
 RA Fritchman J.L., Weidman J.P., Small K.V., Sandosky M., Fuhmann J.L.,
 RA Nguyen D.T., Overbeck R., Weidman J.P., Stuck E.M., Phillips C.A., Merrick J.M.,
 RA Tomb J.-F., Dougherty R.A., Rhee P.W., Hu P.-C., Luster J.S.,
 RA Peterson S.N., Smith H.G., Borodovsky M., Kaine R.P., Venter J.C.;
 RI "The minimal gene complement of Mycoplasma genitalium.";
 RL Science 270:397-403(1995)
 CC -1- FUNCTION: COULD BE A METALLOPROTEASE
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF G-STALOGLYCOPROTEINS; CLEAVES
 CC 31-ARG-1-ASP-32 BOND IN GLYCOPROTEIN A. DOES NOT CLEAVE
 CC UNHYDROXYLATED PROTEINS, HYDROLYZED HYDROXYPROTEINS OF
 CC GLYCOPROTEINS THAT ARE ONLY N-GLYCOSYLATED.
 CC -1- COFACTOR: 71NC (PROBABLY).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M22; ALSO KNOWN AS THE

CC GLYCOPROTEASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.1st-sib.ch/announcements/>
 CC or send an email to 1st-sib.ch@sib.ch).
 CC EMBL: U00684; AAC71262.1; -
 DR TIGR: M0046; -
 DR InterPro: IPR000005; -
 DR Pfam: PF00814; Peptidase_M22; 1.
 DR PRINIS: PR00789; OSIALOPTASE.
 DR PROSITE: PS01016; GLYCOPROTEASE; 1.
 KW Hydrolase; Metalloprotease; zinc.
 FT METAL 110 110 ZINC (POTENTIAL).
 FT METAL 114 114 ZINC (POTENTIAL).
 SQ SEQUENCE 315 AA; 34708 MW; D186ED176C4E574 CRF64;

Query Match 0.54; Score 57; DB 1; Length 315;
 Best Local Similarity 28.84; Pred. No. 1,1e+02;
 Matches 15; Conservative 10; Mismatches 27; Indels 0; Gaps 0;
 QY 575 VLKIAADILKEELISPDGLSNQVVVCLPFVVINNDTSAEMKIAIYLS 626
 Db 256 IREAFEMLLVGVGVANVLSHRIETLN-LFLAISEYTSRCAMIGFTAS 307

RESULT 190
 WZEB_SALTY
 ID WZEB_SALTY STANDARD; PRT; 327 AA.
 AC Q04866;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE CHAIN LENGTH DETERMINANT PROTEIN (POLYSACCHARIDE ANTIGEN CHAIN
 DE REGULATOR).
 GN WZEB OR OLD OR ROL.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria, gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxID=602;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-112;
 RX MEDLINE 93225815; PubMed 7682279;
 RA Hastin D.A., Stevenson G., Brown P.K., Haase A., Reeves P.R.;
 PT "Repeat unit polysaccharides of bacteria: a model for polymerization
 PT resembling that of ribosomes and fatty acid synthetase, with a novel
 PT mechanism for determining chain length.";
 RL Mol. Microbiol. 7:725-734(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STEATH L12;
 RX MEDLINE 92355434; PubMed 1379582;
 RA Batchelor P.A., Allano P., Riffali E., Hull S.L., Hull R.A.;
 FT "Polysaccharide sequences of the group 1 regulatory polysaccharide
 FT antigen chain length (rel) from Escherichia coli and Salmonella
 FT typhimurium: distinct homology and functional complementation.";
 RL J. Bacteriol. 174:5228-5236(1992).
 CC -1- FUNCTION: CONFERS A MODAL DISTRIBUTION OF CHAIN LENGTH ON THE
 CC O-ANTIGEN COMPONENT OF LIPOPOLYSACCHARIDE (LPS). GIVES RISE TO
 CC A REDUCED NUMBER OF SHORT CHAIN MOLECULES AND INCREASES IN
 CC NUMBERS OF LONGER MOLECULES, WITH A MODAL VALUE OF 20.
 CC -1- PATHWAY: LIPOPOLYSACCHARIDE BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: INTRACELLULAR MEMBRANE PROTEIN. INNER MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE WZEB/CML/ROL FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

EMBL: AF010632; AA008323, 1;
 ER
 ER
 HSPF: 100933; 21YS;
 11CR; H12700;
 Indirect: 1700000003;

07 1126 P-109110-VN²PN²AV²SV²*125VAV²VF²EF²FA²CF²GV²Q 1178
Matches 14; Conservative 10; Mismatches 27; Gaps 0;
Best Local Similarity 26.4%; Pred. No. 1.5e+02;

Journal of Management Inquiry 23(4) 391-407

100

